The Philosophical Magazine

FIRST PUBLISHED IN 1798

A Journal of Theoretical Experimental and Applied Physics

Vol. 3 December 1958 No. 36
Eighth Series

£1 5s. 0d., plus postage
Annual Subscription £13 10s. 0d., payable in advance

Printed and Published by
TAYLOR & FRANCIS LTD
RED LION COURT, FLEET STREET, LONDON, E.C.4

THE PHILOSOPHICAL MAGAZINE

Authors wishing to submit papers for publication in the Journal should send manuscripts directly to the Publishers.

Manuscripts should be typed in *double* spacing on one side of quarto (8×10 in.) paper, and authors are urged to aim at absolute clarity of meaning and an attractive presentation of their texts.

References should be listed at the end in alphabetical order of authors and should be cited in the text in terms of author's name and date. Diagrams should normally be in Indian ink on white card, with lettering in soft pencil, the captions being typed on a separate sheet.

A leaflet giving detailed instructions to authors on the preparation of papers is available on request from the Publishers.

Authors are entitled to receive 25 offprints of a paper in the Journal free of charge, and additional offprints can be obtained from the Publishers.

The *Philosophical Magazine* and its companion journal, *Advances in Physics*, will accept papers for publication in experimental and theoretical physics. The *Philosophical Magazine* publishes contributions describing new results, letters to the editor and book reviews. *Advances in Physics* publishes articles surveying the present state of knowledge in any branch of the science in which recent progress has been made. The editors welcome contributions from overseas as well as from the United Kingdom, and papers may be published in English, French and German.

The Initial Stages of Growth of Oriented Copper Nuclei on Single Crystal Surfaces of Silver†

By E. Grunbaum‡, R. C. Newman§ and D. W. Pashley||
Physics Department, Imperial College, London

[Received July 12, 1958]

Abstract

Layers of metallic copper of less than 1 Å in average thickness have been deposited in a controlled manner on to smooth (111) surfaces of silver inside an electron diffraction camera. It is found that an electron diffraction pattern from oriented copper nuclei appears suddenly at some critical thickness. This critical thickness decreases with increasing substrate temperature, being about 0·9 Å at room temperature and 0·2 Å at 300°c. Possibilities for the initial structure of the deposit, and for the change which occurs at the critical thickness, are considered, and the effect of substrate temperature is discussed.

§ 1. Introduction

In a previous paper (Newman and Pashley 1955) results were given for the growth of extremely thin evaporated layers of copper on (111) faces of silver. It was found that, for an average deposit thickness of less than about 0·8 Å, no visible change occurred in electron diffraction reflection patterns from the substrate surface. As the layer was increased in thickness beyond this value, a diffraction pattern due to oriented copper crystallites appeared, and rapidly increased in intensity. It was supposed that this initial appearance of a diffraction pattern from copper was controlled by a limiting sensitivity of detection.

This paper reports the results of a more careful study of the rate of increase in intensity of the spots due to diffraction by copper, carried out in 1955. This work has shown that the pattern appears extremely rapidly, and that the change from no spots due to copper, to quite a strong visible pattern occurs for a thickness increase of only about 0·05 Å. This did not seem to be consistent with a limiting sensitivity, and it was considered that a more likely explanation is that a transition in the mode of growth occurs at a certain stage. In order to investigate the effect further, the following experiments were carried out: (1) the diffraction patterns have been studied over a range of electron energies;

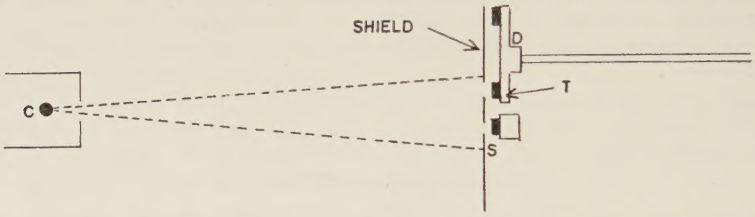
† Communicated by the Authors.
‡ Now at the University of Chile, Santiago.
§ Now at the A.E.I. Research Laboratory, Aldermaston, Berks.
|| Now at the T.I. Research Laboratories, Hinxton Hall, Cambridge.

(2) the effect of varying the substrate temperature has been investigated. The results of both of these experiments, reported below, support this latter view.

§ 2. Experimental Technique

The copper was deposited on (111) faces of silver crystals in position in the electron diffraction camera; radioactive copper was used in order to allow accurate thickness measurements to be made. In the previous work (Newman and Pashley 1955) a large number of specimens of different thicknesses was prepared and examined, in order to investigate thickness effects. As this method is rather laborious, a method was developed to allow a series of measurements to be made on one single specimen which remains *in situ* in the diffraction camera. The automatic recording technique described by Kehoe *et al.* (1956) was used, and the rate of deposition was determined in the following manner (see fig. 1).

Fig. 1



Radioactive copper (C) was deposited simultaneously on to the specimen (S) under examination, and on to a second similar specimen (T) which was one of a number mounted on the disc D. During the experiment, D was rotated at intervals so that each specimen on the disc received a copper deposit for a certain period during the deposition. At the end of the experiment, the thickness of the copper deposit on each specimen such as T was determined from its radioactivity, and a growth curve was plotted. A check on the rate of growth on the specimen (S) was made by means of a determination of the total thickness on this specimen. The technique was, in effect, used to check that the rate of growth was constant, so that accurate interpolations could be made to obtain the thickness of the layer at any given instant at which a diffraction pattern was recorded. In all results reported below, the rate of growth was accurately constant during the entire deposition.

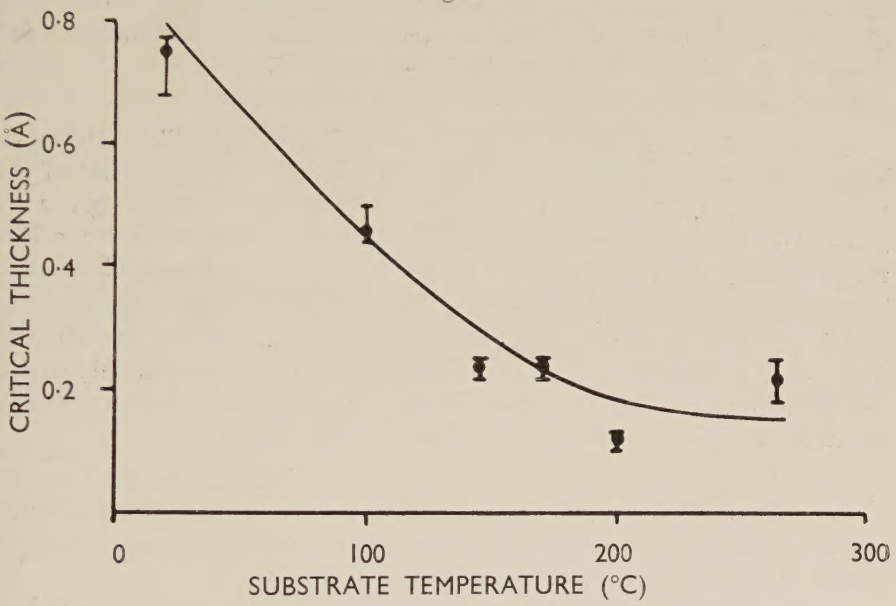
§ 3. Experiments at Various Electron Energies

Several runs were carried out at a number of different accelerating voltages in the range 10–60 kv. In each case, the mean thickness (hereafter called the critical thickness) at which the copper deposit suddenly appeared on the diffraction pattern was determined. This critical thickness was found to be the same in all cases, within experimental accuracy, and to be 0·9 Å, in reasonable agreement with previous results (Newman and Pashley 1955).

§ 4. Experiments on Heated Substrates

The deposition of the copper was studied on substrates which were maintained at temperatures ranging from 20–300°C, by means of the furnace described previously (Kehoe *et al.* 1956). It was found that on all specimens, the pattern due to oriented copper appeared suddenly at some critical thickness, which depended upon the substrate temperature as shown in fig. 2. As the substrate temperature is increased, the critical thickness becomes less, reaching a value of about 0·2 Å at 300°C.

Fig. 2



§ 5. Discussion

It seems unlikely that the sensitivity of detection of the copper deposits should be independent of the electron energy over the range 10–60 kv, since the inelastic and elastic scattering cross sections both vary considerably and differently over this range. This variation causes marked changes in the depth of penetration of the electrons, which would be expected to have a considerable effect on the proportion of the electrons diffracted by any surface film. The first experiment therefore gives strong support to the hypothesis that the critical thickness arises from some other cause than that of a limit of detection.

The second experiment confirms this view, since the critical thickness is found to depend upon temperature. It would be expected that the limit of detection (in terms of average thickness/unit area) would be independent of the substrate temperature, provided the limit of detection is not influenced by changes in the form of the nuclei. If the nuclei are

sufficiently small to allow complete penetration, without any appreciable absorption, the sensitivity of detection will have a maximum value. For nuclei which are sufficiently large to give rise to appreciable absorption effects, the larger the nuclei, the lower the sensitivity of detection. However, in the second experiment, the nuclei formed at the elevated temperatures are the larger, as indicated by the sharpness of the diffraction pattern, and yet they are detected at lower average thicknesses. This proves that the initial appearance of the diffraction patterns due to oriented copper nuclei is not related to a limiting sensitivity of detection, but arises from some property of the growth of the deposit. It is therefore necessary to consider (*a*) what form is taken by the initial deposits, such that they give rise to no observable diffraction patterns, and (*b*) what change takes place at the critical thicknesses indicated in fig. 2.

There are two possibilities which would explain (*a*):

(1) The initial deposit is in the form of small crystalline islands of monatomic thickness which have the same lateral spacing as that of the substrate silver. This is the mechanism for epitaxial growth given by Frank and Van der Merwe (1949 a, b) except that these authors consider the possibility of stability of the nuclei up to much greater thicknesses.

(2) The initial deposit occurs as a distribution of single copper atoms which are able to diffuse over the silver surface. At any instant the position of a particular atom is determined by the substrate only and is not influenced by adjacent copper atoms. The copper atoms may thus be regarded as a two-dimensional gas (see Frenkel 1923).

In both cases, it is reasonable that re-crystallization to give copper nuclei with their normal spacing should occur at some critical average surface density denoted by n_c. The expected dependence of n_c on temperature will now be examined for each case.

In case (1) the pseudomorphic islands will re-crystallize when dislocations are generated in the islands. Since an activation energy is required for this process, re-crystallization is expected to occur more readily at higher temperatures. This temperature dependence will be enhanced by the growth of larger islands at higher temperatures due to the higher surface mobility of the depositing copper atoms; these larger islands would be inherently less stable than small islands as the probability of a dislocation being generated in a particular island is dependent on its size. The experimental observations would therefore be explained qualitatively by this mechanism.

In case (2), it is assumed that the silver substrate exerts a periodic binding force on each copper atom. If the energy barrier between adjacent sites is E, the number of mobile atoms at any instant will be given by the total number present on the surface, multiplied by the appropriate Boltzmann factor. Crystallization will occur when the local surface density exceeds a critical value such that the nucleus is formed and does not immediately re-evaporate. The frequency with which this occurs

will obviously depend on the mobility of the copper atoms. Thus, the value of n_c when crystallization occurs will be given by

$$n_c = N_c \exp\{+(E/kT)\}$$

where N_c is related to the critical number of mobile atoms required to form a nucleus. This mechanism would explain the observations provided N_c does not vary too rapidly with temperature in opposition to the exponential factor.

Unfortunately, it is not possible to calculate the activation energy for the generation of dislocations in pseudomorphic metal layers, as the force laws involved are not known. Also, the energy E above cannot be computed with any accuracy, due to the presence of adsorbed gases on the substrate surface during the deposition. It would thus appear that the experimental evidence is insufficient to allow any distinction to be made between the two possible mechanisms which have been suggested.

The present results must lead to a modification of a conclusion concerning sensitivity which was reported previously (Newman and Pashley 1955). It was stated that the higher sensitivity of electron diffraction for detecting silver bromide (a minimum average layer of 0·4 Å was detected) compared with copper deposits was due to differences in the scattering factors of the elements present in the layers. This is now contradicted by the present results where a copper deposit of only 0·2 Å in mean thickness has been detected. It therefore seems likely that the first appearance of the silver bromide is controlled by the mechanism of growth rather than by the limiting sensitivity of detection.

Ackowledgments

The authors wish to thank Dr. M. Blackman for useful discussions. One of us (E. G.) expresses his gratitude to the British Council for a scholarship and grant, and to The University of Chile for leave of absence. Another of us (R. C. N.) is indebted to the Department of Scientific and Industrial Research for a maintenance grant. The work was carried out during the tenure of an Imperial Chemical Industries Research Fellowship by the third author.

References

Frank, F. C., and Van Der Merwe, J. H., 1949 a, *Proc. roy. Soc.* A, **198**, 205; 1949 b, *Ibid.*, **200**, 125.

Frenkel, J., 1923, *Z. Phys.*, **26**, 117.

Kehoe, R. B., Newman, R. C., and Pashley, D. W., 1956, *Brit. J. appl. Phys.*, **7**, 29.

Newman, R. C., and Pashley, D. W., 1955, *Phil. Mag.*, **46**, 927.

On the Thermal Instability of a Rotating Fluid Sphere†

By F. E. Bisshopp
Department of Physics and Yerkes Observatory, The University of Chicago

[Received July 14, 1958]

Summary

In this paper the problem of the thermal instability of a self-gravitating rotating fluid sphere of constant density with internal heating is treated. A method of solving this problem is described which correctly takes into account the full set of boundary conditions. The cases when the bounding sphere is rigid and when it is free are both treated. The Rayleigh–Ritz method of approximating lowest characteristic values is used in a variation principle presented by Chandrasekhar to calculate the critical Rayleigh number as a function of the Taylor number.

§ 1. Introduction

In a recent paper Chandrasekhar has studied the problem of the onset of thermal instability in a rotating fluid sphere heated within (Chandrasekhar 1957 a) and has determined the critical Rayleigh number as a function of the Taylor number for a specific set of boundary conditions which corresponds neither to a rigid boundary nor to a free one at the surface of the sphere. The aim of the present paper is to re-examine the same problem with the correct boundary conditions.

In the formulation of the problem, axial symmetry of the perturbed motions and of the modified temperature distribution is assumed. The assumed symmetry allows the expression of the velocity perturbation as the superposition of a poloidal and a toroidal field in terms of two azimuth-independent scalars. As we shall see in § 2, the equations connecting these two scalars and the function which characterizes the temperature perturbation are a set of three partial differential equations, two of order two and one of order four. In these equations the Taylor number, **T**, can be considered a parameter to be specified in advance, in which case the equations together with the boundary conditions determine a characteristic value problem for the Rayleigh number, **R**. The critical Rayleigh number for the onset of convection is to be found by evaluating the lowest characteristic value. An entirely equivalent procedure which we will have no need of considering is to consider the Rayleigh number as the parameter and the Taylor number as the characteristic value.

When one is confronted with a problem of the type described above, one might, in general, expect to proceed in either of the following two ways: Upon obtaining the most general solution of the differential equations and

† Communicated by S. Chandrasekhar, F.R.S.

requiring that the solution satisfy the boundary conditions, one obtains an equation to determine the set of characteristic values; *or*, upon obtaining the most general set of functions which satisfy the boundary conditions and requiring that the solution satisfy the differential equations, one again obtains an equation to determine the set of characteristic values. Because of the complexity of the problem we are considering, it is necessary to employ a method which contains elements of both the above mentioned procedures. We shall first construct a set of functions which satisfy all the boundary conditions and the two second order differential equations. By requiring that these functions satisfy the remaining fourth order equation (or the equivalent variational equation) we shall determine the critical Rayleigh number as a function of the parameter $\mathbf{T}$.

In constructing the set of trial functions for the variational equation we will find it necessary to expand the defining scalar of the poloidal field in a complete set of functions which satisfy *two* boundary conditions at $r=1$ and which are bounded at $r=0$. A discussion of similar expansions can be found in a paper by Chandrasekhar and Reid (1957).

§ 2. Equations of the Problem

In Chandrasekhar's paper (1957 a) the equations governing marginal stability in the case where rigid rotation is present are presented and discussed adequately. We shall be content with merely presenting them here.

Let us consider a homogeneous fluid sphere of radius r_0, rotating with a constant angular velocity $\mathbf{\Omega}$, about the z-axis. We assume there is within the sphere a uniform distribution of heat source which, in the absence of conduction, will cause the temperature at every point to rise at the rate ϵ. In addition we assume no external forces acting on the sphere, we ignore rotational flattening, we assume axial symmetry of any perturbations, and we take into account the variation of density due to thermal expansion in the Boussinesque approximation.

The equilibrium temperature distribution is given by (Chandrasekhar 1952, 1957 a)

$$T=\beta(r_0^2-r^2); \quad \beta=\frac{\epsilon}{6\varkappa}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where $\varkappa$ is the coefficient of thermal conductivity.

The expression of the perturbed velocity field as the superposition of a poloidal and a toroidal field in terms of the two azimuth-independent scalars, U and V, can be written as follows (Chandrasekhar 1956 a, b, 1957 a, Lüst and Schlüter 1954):

$$\mathbf{u}=\mathbf{l}_z\times\mathbf{r}\,V(r,\mu)+\operatorname{curl}\,[\mathbf{l}_z\times\mathbf{r}U(r,\mu)]. \quad . \quad . \quad . \quad . \quad (2)$$

In terms of spherical coordinates $(r,\ \mu=\cos\vartheta,\ \varphi)$,

$$\mathbf{u}=-\frac{\partial}{\partial\mu}[(1-\mu^2)U]\mathbf{l}_r-\frac{\sqrt{(1-\mu^2)}}{r}\frac{\partial}{\partial r}[r^2U]\mathbf{l}_\vartheta+r\sqrt{(1-\mu^2)}V\mathbf{l}_\varphi. \quad (3)$$

The equations governing marginal stability assume the following form:

$$\Delta_5^2 U - \frac{2\Omega}{\nu}\frac{\partial V}{\partial z} = \frac{\gamma}{\nu r}\frac{\partial \theta}{\partial \mu}, \quad \ldots\ldots (4)$$

$$\Delta_5 V + \frac{2\Omega}{\nu}\frac{\partial U}{\partial z} = 0, \quad \ldots\ldots (5)$$

$$\Delta_3 \theta = \frac{2\beta}{\kappa} r \frac{\partial}{\partial \mu}[(1-\mu^2)U], \quad \ldots\ldots (6)$$

where θ is the deviation from the equilibrium temperature distribution, ν is the coefficient of kinematic viscosity, $\gamma = 4\pi G\alpha\rho/3$, G is the gravitation constant, α is the coefficient of volume expansion, ρ is the density,

$$\Delta_5 = \frac{\partial^2}{\partial r^2} + \frac{4}{r}\frac{\partial}{\partial r} + \frac{(1-\mu^2)}{r^2}\frac{\partial^2}{\partial \mu^2} - \frac{4\mu}{r^2}\frac{\partial}{\partial \mu}, \quad \ldots\ldots (7)$$

and

$$\Delta_3 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{(1-\mu^2)}{r^2}\frac{\partial^2}{\partial \mu^2} - \frac{2\mu}{r^2}\frac{\partial}{\partial \mu}, \quad \ldots\ldots (8)$$

are the five- and three-dimensional Laplacian operators for axisymmetric functions.

By the transformation to the dimensionless variables,

$$\left.\begin{aligned} \mathbf{r}' &= \mathbf{r}/r_0, \\ U'(r', \mu) &= U(r, \mu)/r_0\Omega, \\ V'(r', \mu) &= \nu V(r, \mu)/2{r_0}^2\Omega^2, \\ \theta'(r', \mu) &= \kappa\theta(r, \mu)/2\beta{r_0}^4\Omega, \end{aligned}\right\} \quad \ldots\ldots (9)$$

the equations assume the more convenient forms,

$$\Delta_5^2 U - \mathbf{T}\frac{\partial V}{\partial z} = \mathbf{R}\frac{1}{r}\frac{\partial \theta}{\partial \mu}, \quad \ldots\ldots (10)$$

$$\Delta_5 V + \frac{\partial U}{\partial z} = 0, \quad \ldots\ldots (11)$$

and

$$\Delta_3 \theta = r\frac{\partial}{\partial \mu}[(1-\mu^2)U], \quad \ldots\ldots (12)$$

where

$$\mathbf{R} = \frac{2\beta\gamma}{\kappa\nu}{r_0}^6; \quad \mathbf{T} = \frac{4\Omega^2}{\nu^2}{r_0}^4, \quad \ldots\ldots (13)$$

and the primes have been suppressed. The number denoted $\mathbf{R}$ is the Rayleigh number; $\mathbf{T}$ is the Taylor number.

We now make use of the easily verified identity,

$$\frac{1}{r}\frac{\partial}{\partial \mu}\Delta_3 = \Delta_5\frac{1}{r}\frac{\partial}{\partial \mu} \quad \ldots\ldots (14)$$

to rewrite eqn. (12) as

$$\Delta_5\left(\frac{1}{r}\frac{\partial \theta}{\partial \mu}\right) = \frac{\partial^2}{\partial \mu^2}[(1-\mu^2)U]. \quad \ldots\ldots (15)$$

The boundary conditions which must be satisfied by the solutions of eqns. (10), (11) and (15) depend upon whether the bounding surface at $r=1$ is assumed to be a rigid or a free boundary. In both cases we require the vanishing of the radial component of the velocity and of the deviation of the temperature from the equilibrium value. At a rigid boundary we require the vanishing of u_ϑ and u_φ and at a free boundary we require the vanishing of the tangential viscous stresses. Accordingly, the boundary conditions for a rigid boundary at $r=1$ can be written;

$$U=0,\ \frac{\partial U}{\partial r}=0,\ \frac{\partial \theta}{\partial \mu}=0,\ V=0;\quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

and for a free boundary at $r=1$;

$$U=0,\ \frac{\partial^2}{\partial r^2}(rU)=0,\ \frac{\partial \theta}{\partial \mu}=0,\ \frac{\partial V}{\partial r}=0. \quad . \quad . \quad . \quad . \quad . \quad (17)$$

§ 3. The Variational Principle

Chandrasekhar has shown that eqns. (10), (11) and (12) together with either of the sets of boundary conditions (16) and (17) imply the following formula which provides the basis of a variational method for determining $\mathbf{R}$ in terms of the parameter $\mathbf{T}$ (Chandrasekhar 1957 b),

$$\mathbf{R}\int_0^1 r^2\,dr\int_{-1}^1 d\mu\,|\Delta\theta|^2=-2\int_{-1}^1(1-\mu^2)\,d\mu\left(\frac{\partial U}{\partial r}\right)^2_{r=1}$$
$$+\int_0^1 r^4\,dr\int_{-1}^1(1-\mu^2)\,d\mu\left\{(\Delta_5 U)^2+\mathbf{T}\left[\left(\frac{\partial V}{\partial r}\right)^2+\frac{(1-\mu^2)}{r^2}\left(\frac{\partial V}{\partial \mu}\right)^2\right]\right\}. \quad . \quad . \quad . \quad (18)$$

For reasons of convenience we will use the following relation from which eqn. (18) is derived:

$$\mathbf{R}=\frac{\int_0^1 r^4\,dr\int_{-1}^1(1-\mu^2)\,d\mu[U\Delta_5{}^2U-\mathbf{T}U(\partial V/\partial z)]}{\int_0^1 r^4\,dr\int_{-1}^1(1-\mu^2)\,d\mu U(\partial\theta/r\partial\mu)}. \quad . \quad . \quad (19)$$

The application of the Rayleigh–Ritz method consists in substituting an appropriate trial function, dependent on several arbitrary parameters, in eqn. (19) and minimizing with respect to the parameters. The requirements on the functions substituted for U, V and θ in eqn. (19) are, (1) that they satisfy the correct boundary conditions and (2) that they be consistent with eqns. (11) and (15). Keeping these requirements in mind, we now proceed to develop an appropriate set of functions.

§ 4. The Expansion of the Poloidal Field

We begin the construction of the trial functions by expanding the poloidal scalar, U, in terms of the complete set of functions determined by the equation

$$\Delta_5{}^2U=\alpha^4U \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (20)$$

where Δ_5 is again the five-dimensional Laplacian for axisymmetric functions. We find that by using this particular set of functions we are able to solve eqns. (11) and (15) with little effort and that, in addition, we are able to evaluate the integrals encountered in eqn. (19) analytically.

The general solution of eqn. (20) which is bounded at the origin is

$$U=\sum_{n=0}^{\infty}[A_nJ_{n+3/2}(\alpha r)+B_nI_{n+3/2}(\alpha r)]C_n^{3/2}(\mu)/r^{3/2}, \quad . \quad . \quad (21)$$

where $C_n^{3/2}$ are the Gegenbauer polynomials and $J_{n+3/2}$ and $I_{n+3/2}$ are the Bessel functions of half-odd-integral index for real and imaginary argument (Sommerfeld 1949, Erdelyi *et al.* 1953).

The application of the boundary condition $U(1, \mu)=0$ implies

$$U=\sum_{n=0}^{\infty}A_n\left[\frac{J_{n+3/2}(\alpha r)}{J_{n+3/2}(\alpha)}-\frac{I_{n+3/2}(\alpha r)}{I_{n+3/2}(\alpha)}\right]C_n^{3/2}(\mu)/r^{3/2}. \quad . \quad . \quad (22)$$

On applying the second boundary condition at $r=1$, we obtain for the case of a rigid boundary,

$$\frac{J'_{n+3/2}(\alpha)}{J_{n+3/2}(\alpha)}=\frac{I'_{n+3/2}(\alpha)}{I_{n+3/2}(\alpha)}; \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

and for a free boundary,

$$\frac{J'_{n+3/2}(\alpha)}{J_{n+3/2}(\alpha)}=\frac{I'_{n+3/2}(\alpha)}{I_{n+3/2}(\alpha)}-\alpha. \quad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

Thus we may write

$$U=\sum_{n=0}^{\infty}\sum_{j=1}^{\infty}A_{nj}\left[\frac{J_{n+3/2}(\alpha_{nj}r)}{J_{n+3/2}(\alpha_{nj})}-\frac{I_{n+3/2}(\alpha_{nj}r)}{I_{n+3/2}(\alpha_{nj})}\right]C_n^{3/2}(\mu)/r^{3/2}, \quad (25)$$

where α_{nj} is the jth root of eqn. (23) when the boundary at $r=1$ is rigid and of eqn. (24) when the boundary is free.

§ 5. The Temperature Perturbation and the Toroidal Field

We will find it convenient to introduce the following notation before solving eqns. (11) and (15):

$$f(\alpha_{nj}r)\equiv f_{n/n}(\alpha_{nj}r)\equiv\frac{J_{n+3/2}(\alpha_{nj}r)}{J_{n+3/2}(\alpha_{nj})}-\frac{I_{n+3/2}(\alpha_{nj}r)}{I_{n+3/2}(\alpha_{nj})}, \quad . \quad . \quad . \quad (26)$$

$$g(\alpha_{nj}r)\equiv g_{n/n}(\alpha_{nj}r)\equiv\frac{J_{n+3/2}(\alpha_{nj}r)}{J_{n+3/2}(\alpha_{nj})}+\frac{I_{n+3/2}(\alpha_{nj}r)}{I_{n+3/2}(\alpha_{nj})}, \quad . \quad . \quad . \quad (27)$$

$$f_{l/m}(\alpha_{nj}r)\equiv\frac{J_{l+3/2}(\alpha_{nj}r)}{J_{m+3/2}(\alpha_{mj})}-\frac{I_{l+3/2}(\alpha_{nj}r)}{I_{m+3/2}(\alpha_{mj})}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (28)$$

$$g_{l/m}(\alpha_{nj}r)\equiv\frac{J_{l+3/2}(\alpha_{nj}r)}{J_{m+3/2}(\alpha_{mi})}+\frac{I_{l+3/2}(\alpha_{nj}r)}{I_{m+3/2}(\alpha_{mj})}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (29)$$

The following relations are easily verified:

$$\Delta_5[f_{m/n}(\alpha_{nj}r)C_m^{3/2}(\mu)/r^{3/2}] = -\alpha_{nj}^2 g_{m/n}(\alpha_{nj}r)C_m^{3/2}(\mu)/r^{3/2}, \quad (30)$$

$$\Delta_5[g_{m/n}(\alpha_{nj}r)C_m^{3/2}(\mu)/r^{3/2}] = -\alpha_{nj}^2 f_{m/n}(\alpha_{nj}r)C_m^{3/2}(\mu)/r^{3/2}, \quad (31)$$

$$\frac{\partial^2}{\partial\mu^2}[(1-\mu^2)C_n^{3/2}(\mu)] = -(n+1)(n+2)C_n^{3/2}(\mu), \quad . \quad . \quad . \quad (32)$$

$$\frac{\partial}{\partial z}\left[f_{m/n}(\alpha_{nj}r)\frac{C_m^{3/2}(\mu)}{r^{3/2}}\right] = \frac{\alpha_{nj}}{r^{3/2}}\left[\left(\frac{m+2}{2m+3}\right)f_{m-1/n}(\alpha_{nj}r)C_{m-1}^{3/2}(\mu) - \left(\frac{m+1}{2m+3}\right)g_{m+1/n}(\alpha_{nj}r)C_{m+1}^{3/2}(\mu)\right]; \; m \neq 0, \quad . \quad . \quad (33)$$

$$\frac{\partial}{\partial z}\left[g_{m/n}(\alpha_{nj}r)\frac{C_m^{3/2}(\mu)}{r^{3/2}}\right] = \frac{\alpha_{nj}}{r^{3/2}}\left[\left(\frac{m+2}{2m+3}\right)g_{m-1/n}(\alpha_{nj}r)C_{m-1}^{3/2}(\mu) - \left(\frac{m+1}{2m+3}\right)f_{m+1/n}(\alpha_{nj}r)C_{m+1}^{3/2}(\mu)\right]; \; m \neq 0, \quad . \quad . \quad (34)$$

$$\frac{\partial}{\partial z}\left[f_{0/n}(\alpha_{nj}r)\frac{C_0^{3/2}(\mu)}{r^{3/2}}\right] = -\frac{\alpha_{nj}}{r^{3/2}}\left(\frac{1}{3}\right)g_{1/n}(\alpha_{nj}r)C_1^{3/2}(\mu), \quad . \quad . \quad (35)$$

$$\frac{\partial}{\partial z}\left[g_{0/n}(\alpha_{nj}r)\frac{C_0^{3/2}(\mu)}{r^{3/2}}\right] = -\frac{\alpha_{nj}}{r^{3/2}}\left(\frac{1}{3}\right)f_{1/n}(\alpha_{nj}r)C_1^{3/2}(\mu), \quad . \quad . \quad (36)$$

$$\frac{\partial}{\partial z}[r^n C_n^{3/2}(\mu)] = (n+2)r^{n-1}C_{n-1}^{3/2}(\mu); \; n \neq 0, \quad . \quad . \quad . \quad . \quad . \quad . \quad (37)$$

$$\frac{\partial}{\partial z}[C_0^{3/2}(\mu)] = 0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)$$

Equations (30 and (31) follow from the differential eqn. (20) and eqns. (32) through (38) are derived with the help of the recursion relations for the Bessel functions and the Gegenbauer polynomials and the further relation

$$\frac{\partial}{\partial z} = \mu\frac{\partial}{\partial r} + \frac{(1-\mu^2)}{r}\frac{\partial}{\partial\mu}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (39)$$

Armed with eqns. (30) through (38), we are now prepared to attack eqns. (11) and (15).

We begin by expanding $\partial\theta/r\partial\mu$ in terms of the Gegenbauer polynomials,

$$\frac{1}{r}\frac{\partial\theta}{\partial\mu} = \sum_{n=0}^{\infty} T_n(r)C_n^{3/2}(\mu). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (40)$$

Using eqns. (31) and (32) we obtain a particular solution of eqn. (15) to which must be added the bounded solution of

$$\Delta_5\left(\frac{1}{r}\frac{\partial\theta}{\partial\mu}\right) = 0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (41)$$

The result is the following expression for $T_n(r)$:

$$T_n(r) = \sum_{j=1}^{\infty}\frac{(n+1)(n+2)}{\alpha_{nj}^2}\frac{A_{nj}}{r^{3/2}}[g(\alpha_{nj}r) + B_{nj}r^{n+3/2}]. \quad . \quad . \quad (42)$$

The application of the boundary condition $\partial\theta/\partial\mu = 0$ at $r = 1$ implies

$$B_{nj} = -g(\alpha_{nj}) = -2. \quad \ldots \quad (43)$$

Equation (40) becomes finally

$$\frac{1}{r}\frac{\partial\theta}{\partial\mu} = \sum_{n=0}^{\infty}\sum_{j=1}^{\infty} A_{nj}\frac{(n+1)(n+2)}{{\alpha_{nj}}^2}\left[g(\alpha_{nj}r) - 2r^{n+3/2}\right]\frac{{C_n}^{3/2}(\mu)}{r^{3/2}}. \quad (44)$$

The determination of V is only slightly more complicated than the determination of θ. We begin by writing, with the help of eqns. (33) and (35), the expression for $\partial U/\partial z$.

$$\frac{\partial U}{\partial z} = \sum_{n=0}^{\infty}\sum_{j=1}^{\infty}\frac{{C_n}^{3/2}(\mu)}{r^{3/2}}\left[A_{n+1,j}\alpha_{n+1,j}\left(\frac{n+3}{2n+5}\right)f_{n/n+1}(\alpha_{n+1,j}r) - A_{n-1,j}\alpha_{n-1,j}\left(\frac{n}{2n+1}\right)g_{n/n-1}(\alpha_{n-1,j}r)\right]. \quad (45)$$

Using this expression in eqn. (11), we obtain for V:

$$V = \sum_{n=0}^{\infty}\sum_{j=1}^{\infty}\frac{{C_n}^{3/2}(\mu)}{r^{3/2}}\left[\frac{A_{n+1,j}}{\alpha_{n+1,j}}\left(\frac{n+3}{2n+5}\right)g_{n/n+1}(\alpha_{n+1,j}r) - \frac{A_{n-1,j}}{\alpha_{n-1,j}}\left(\frac{n}{2n+1}\right)f_{n/n-1}(\alpha_{n-1,j}r) + B_{nj}r^{n+3/2}\right] \quad (46)$$

where the constant B_{nj} is to be determined by the boundary condition on V at $r = 1$. For a rigid boundary, one obtains

$$\frac{A_{n+1,j}}{\alpha_{n+1,j}}\left(\frac{n+3}{2n+5}\right)g_{n/n+1}(\alpha_{n+1,j}) - \frac{A_{n-1,j}}{\alpha_{n-1,j}}\left(\frac{n}{2n+1}\right)f_{n/n-1}(\alpha_{n-1,j}) + B_{nj} = 0, \quad \ldots \quad (47)$$

and for a free boundary,

$$\frac{A_{n+1,j}}{\alpha_{n+1,j}}\left(\frac{n+3}{2n+5}\right)\left[\alpha_{n+1,j}\,g'_{n/n+1}(\alpha_{n-1,j}) - \tfrac{3}{2}g_{n/n+1}(\alpha_{n+1,j})\right] - \frac{A_{n-1,j}}{\alpha_{n-1,j}}\left(\frac{n}{2n+1}\right)\left[\alpha_{n-1,j}f'_{n/n-1}(\alpha_{n-1,j}) - \tfrac{3}{2}f_{n/n-1}(\alpha_{n-1,j})\right] + nB_{nj} = 0, \quad (48)$$

where we have made the following additional definitions:

$$f'_{n/m}(\alpha_{mj}) \equiv \frac{J'_{n+3/2}(\alpha_{mj})}{J_{m+3/2}(\alpha_{mj})} - \frac{I'_{n+3/2}(\alpha_{mj})}{I_{m+3/2}(\alpha_{mj})}, \quad \ldots \quad (49)$$

$$g'_{n/m}(\alpha_{mj}) \equiv \frac{J'_{n+3/2}(\alpha_{mj})}{J_{m+3/2}(\alpha_{mj})} + \frac{I'_{n+3/2}(\alpha_{mj})}{I_{m+3/2}(\alpha_{mj})}. \quad \ldots \quad (50)$$

Equations (47) and (48), together with the recursion relations for Bessel functions, lead to the following expression for a rigid boundary;

$$V = \sum_{n=0}^{\infty}\sum_{j=1}^{\infty}\left\{\frac{A_{n+1,j}}{\alpha_{n+1,j}}\left(\frac{n+3}{2n+5}\right)\left[g_{n/n+1}(\alpha_{n+1,j}r) - g_{n/n+1}(\alpha_{n+1,j})r^{n+3/2}\right] - \frac{A_{n-1,j}}{\alpha_{n-1,j}}\left(\frac{n}{2n+1}\right)\left[f_{n/n-1}(\alpha_{n-1,j}r) - f_{n/n-1}(\alpha_{n-1,j})r^{n+3/2}\right]\right\}\frac{{C_n}^{3/2}(\mu)}{r^{3/2}}, \quad (51)$$

and for a free boundary,

$$V=\sum_{j=1}^{\infty}\left[\frac{A_{1j}}{\alpha_{1j}}(\tfrac{3}{5})g_{0/1}(\alpha_{1j}r)-B_{0j}r^{3/2}\right]\frac{C_0^{3/2}(\mu)}{r^{3/2}}$$
$$+\sum_{n=1}^{\infty}\sum_{j=1}^{\infty}\left\{\frac{A_{n+1,j}}{\alpha_{n+1,j}}\left(\frac{n+3}{2n+5}\right)[g_{n/n+1}(\alpha_{n+1,j}r)-g_{n/n+1}(\alpha_{n+1,j})r^{n+3/2}]\right.$$
$$\left.-\frac{A_{n-1,j}}{\alpha_{n-1,j}}\left(\frac{n}{2n+1}\right)\left[f_{n\,n-1}(\alpha_{n-1,j}r)+\left(\frac{n+3}{n}\right)f_{n/n-1}(\alpha_{n-1,j})r^{n+3/2}\right]\right\}$$
$$\times\frac{C_n^{3/2}(\mu)}{r^{3/2}}. \qquad (52)$$

In particular in deriving eqn. (52) we have made use of the following two relations

$$\alpha_{n+1,j}\,g'_{n/n+1}(\alpha_{n+1,j})-\tfrac{3}{2}g_{n/n+1}(\alpha_{n+1,j})=ng_{n/n+1}(\alpha_{n+1,j})-f(\alpha_{n+1,j}), \qquad (53)$$

$$\alpha_{n-1,j}f'_{n/n-1}(\alpha_{n-1,j})-\tfrac{3}{2}f_{n/n-1}(\alpha_{n-1,j})=-(n+3)f_{n/n-1}(\alpha_{n-1,j})+f(\alpha_{n-1,j}). \qquad (54)$$

Operating on eqns. (51) and (52) with $\partial/\partial z$ we obtain

$$\frac{\partial V}{\partial z}=\sum_{n=0}^{\infty}\sum_{j=1}^{\infty}\left\{A_{n+2,j}\left(\frac{n+4}{2n+7}\right)\left[\left(\frac{n+3}{2n+5}\right)g_{n/n+2}(\alpha_{n+2,j}r)\right.\right.$$
$$\left.-\frac{(n+3)}{\alpha_{n+2,j}}g_{n+1/n+2}(\alpha_{n+2,j})r^{n+3/2}\right]$$
$$-A_{nj}\left[\left(\frac{n+1}{2n+3}\right)\left(\frac{n+3}{2n+5}\right)+\left(\frac{n}{2n+1}\right)\left(\frac{n+2}{2n+3}\right)\right]f(\alpha_{nj}r)$$
$$+p\frac{A_{nj}}{\alpha_{nj}}\left(\frac{n+3}{2n+3}\right)f_{n+1/n}(\alpha_{nj})r^{n+3/2}$$
$$\left.+A_{n-2,j}\left(\frac{n-1}{2n-1}\right)\left(\frac{n}{2n+1}\right)g_{n/n-2}(\alpha_{n-2,j}r)\right\}\frac{C_n^{3/2}(\mu)}{r^{3/2}}. \qquad (55)$$

where

$$\left.\begin{aligned}&p=-(n+1)\text{ for a rigid boundary}\\&p=(n+4)\quad\text{for a free boundary}\end{aligned}\right\}. \qquad (56)$$

§ 6. Application of the Rayleigh–Ritz Method

In order to make the minimization of the functional (19) with respect to the expansion coefficients A_{nj} more tractable, we introduce the following shorter notation:

$$U=\sum_{n=0}^{\infty}\sum_{j=1}^{\infty}A_{nj}F_{nj}(r)C_n^{3/2}(\mu)/r^{3/2}, \qquad (57)$$

$$\frac{1}{r}\frac{\partial\theta}{\partial\mu}=\sum_{n=0}^{\infty}\sum_{j=1}^{\infty}A_{nj}T_{nj}(r)C_n^{3/2}(\mu)/r^{3/2}, \qquad (58)$$

$$\frac{\partial V}{\partial z}=\sum_{n=0}^{\infty}\sum_{j=1}^{\infty}[A_{n-2,j}P_{nj}(r)+A_{nj}Q_{nj}(r)+A_{n+2,j}R_{nj}(r)]\frac{C_n^{3/2}(\mu)}{r^{2/3}}. \qquad (59)$$

Upon differentiating eqn. (19) with respect to A_{mk} one obtains the following equation into which must be substituted the above expressions:

$$\frac{\partial}{\partial A_{mk}}\left\{\int_0^1 r^4\,dr\int_{-1}^1(1-\mu^2)\,d\mu\left[U\Delta_5{}^2U-\mathbf{T}U\frac{\partial V}{\partial z}-\mathbf{R}U\frac{1}{r}\frac{\partial\theta}{\partial\mu}\right]\right\}=0. \qquad (60)$$

Using the orthogonality relation for the Gegenbauer polynomials

$$\int_{-1}^1(1-\mu^2)\,d\mu C_n{}^{3/2}(\mu)C_m{}^{3/2}(\mu)=\frac{2(n+1)(n+2)}{2n+3}\delta_{mn}, \quad . \quad . \quad (61)$$

we obtain

$$\frac{\partial}{\partial A_{mk}}\left\{\sum_{n=0}^{\infty}\sum_{j=1}^{\infty}\sum_{l=1}^{\infty}A_{nj}A_{nl}\alpha_{nl}{}^4\langle nj|F^2|nl\rangle-\mathbf{R}A_{nj}A_{nl}\langle nj|FT|nl\rangle\right.$$
$$-\mathbf{T}A_{nj}[A_{n-2,l}\langle nj|FP|n-2,l\rangle+A_{nl}\langle nj|FQ|nl\rangle$$
$$\left.+A_{n+2,l}\langle nj|FR|n+2,l\rangle]\right\}=0. \quad . \quad . \quad . \quad . \quad . \quad (62)$$

where

$$\langle nj|F^2|nl\rangle = \frac{2(n+1)(n+2)}{2n+3}\int_0^1 r\,dr F_{nj}(r)F_{nl}(r), \qquad (63)$$

$$\langle nj|FT|nl\rangle = \frac{2(n+1)(n+2)}{2n+3}\int_0^1 r\,dr F_{nj}(r)T_{nl}(r), \qquad (64)$$

$$\langle nj|FP|n-2,l\rangle = \frac{2(n+1)(n+2)}{2n+3}\int_0^1 r\,dr F_{nj}(r)P_{nl}(r), \qquad (65)$$

$$\langle nj|FQ|nl\rangle = \frac{2(n+1)(n+2)}{2n+3}\int_0^1 r\,dr F_{nj}(r)Q_{nl}(r), \qquad (66)$$

$$\langle nj|FR|n+2,l\rangle = \frac{2(n+1)(n+2)}{2n+3}\int_0^1 r\,dr F_{nj}(r)R_{nl}(r). \qquad (67)$$

Performing the differentiation indicated in eqn. (62) we obtain

$$\sum_{j=1}^{\infty}A_{mj}\{[\alpha_{mk}{}^4\langle mj|F^2|mk\rangle+\alpha_{mj}{}^4\langle mk|F^2|mj\rangle]$$
$$-\mathbf{R}[\langle mj|FT|mk\rangle+\langle mk|FT|mj\rangle]$$
$$-\mathbf{T}[\langle mj|FQ|mk\rangle+\langle mk|FQ|mj\rangle]\}$$
$$-A_{m-2,j}\mathbf{T}\{\langle mk|FP|m-2,j\rangle+\langle m-2,j|FR|mk\rangle\}$$
$$-A_{m+2,j}\mathbf{T}\{\langle mk|FR|m+2,j\rangle+\langle m+2,j|FP|mk\rangle\}=0. \qquad (68)$$

From the self-adjoint character of eqn. (10) or from the expressions to be derived in § 7 of this paper one can demonstrate the following symmetries of the matrix elements

$$\alpha_{mk}{}^4\langle mj|F^2|mk\rangle = \alpha_{mj}{}^4\langle mk|F^2|mj\rangle, \quad . \quad . \quad . \quad . \quad . \quad . \quad (69)$$

$$\langle mj|FT|mk\rangle = \langle mk|FT|mj\rangle, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (70)$$

$$\langle mj|FQ|mk\rangle = \langle mk|FQ|mj\rangle, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (71)$$

$$\langle mk|FP|m-2,j\rangle = \langle m-2,j|FR|mk\rangle. \quad . \quad . \quad . \quad . \quad (72)$$

To determine the solution of the infinite set of eqns. (68), we set the determinant of the coefficients of A_{mj} equal to zero. The resulting polynomial of infinite degree in $\mathbf{R}$ is

$$|\delta_{mn}[\alpha_{mk}{}^4\langle mj|F^2|mk\rangle - \mathbf{R}\langle mj|FT|mk\rangle - \mathbf{T}\langle mj|FQ|mk\rangle] - \delta_{m-2,n}\mathbf{T}\langle mk|FP|nj\rangle - \delta_{m+2,n}\mathbf{T}\langle mk|FR|nj\rangle| = 0. \quad (73)$$

The critical Rayleigh number is now determined by approximating the above equation with a finite number of terms and computing the lowest root numerically. As may be easily demonstrated, we need only compute matrix elements with even values of m since functions with the subscript m odd do not lead to conservation of angular momentum.

§ 7. Evaluation of the Matrix Elements

The first integral we wish to evaluate is

$$\langle nj|F^2|nl\rangle = \frac{2(n+1)(n+2)}{2n+3}\int_0^1 r\,dr f(\alpha_{nj}r)f(\alpha_{nl}r). \quad . \quad . \quad (74)$$

Using the relations

$$\Delta_5\left[f(\alpha_{nj}r)\frac{C_n{}^{3/2}(\mu)}{r^{3/2}}\right] = \frac{C_n{}^{3/2}(\mu)}{r^{3/2}}\left[\frac{d^2}{dr^2}+\frac{1}{r}\frac{d}{dr}-\frac{(n+\frac{3}{2})^2}{r^2}\right]f(\alpha_{nj}r) \quad (75)$$

and

$$\left[\frac{d^2}{dr^2}+\frac{1}{r}\frac{d}{dr}-\frac{(n+\frac{3}{2})^2}{r^2}\right]f(\alpha_{nj}r) = -\alpha_{nj}{}^2 g(\alpha_{nj}r), \quad . \quad . \quad (76)$$

we can readily show

$$\frac{2n+3}{2(n+1)(n+2)}(\alpha_{nk}{}^4-\alpha_{nj}{}^4)\langle nk|F^2|nj\rangle = \alpha_{nk}{}^2[\alpha_{nj}f'(\alpha_{nj})g(\alpha_{nk})-\alpha_{nk}f(\alpha_{nj})g'(\alpha_{nk})] - \alpha_{nj}{}^2[\alpha_{nk}f'(\alpha_{nk})g(\alpha_{nj})-\alpha_{nj}f(\alpha_{nk})g'(\alpha_{nj})]. \quad (77)$$

Now we use the boundary conditions at $r=1$,

$f(\alpha_{nj})=0$, $f'(\alpha_{nj})=0$, for a rigid boundary;

$f(\alpha_{nj})=0$, $f'(\alpha_{nj})=-\alpha_{nj}$, for a free boundary;

to show

$$\langle nk|F^2|nj\rangle = 0, \quad \alpha_{nj}\neq\alpha_{nk}. \quad . \quad . \quad . \quad . \quad (78)$$

In case $\alpha_{nj}=\alpha_{nk}$ we use l'Hospital's rule to show that for a rigid boundary

$$\langle nk|F^2|nk\rangle=\frac{2(n+1)(n+2)}{2n+3}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (79)$$

and for a free boundary

$$\langle nk|F^2|nk\rangle=-\frac{(n+1)(n+2)}{2n+3}\alpha_{nk}g'(\alpha_{nk}). \quad . \quad . \quad . \quad (80)$$

Using the same techniques, we obtain for either set of boundary conditions

$$\int_0^1 r\,drf(\alpha_{nk}r)g(\alpha_{nj}r)=\frac{2{\alpha_{nk}}^2}{{\alpha_{nk}}^4-{\alpha_{nj}}^4}[\alpha_{nj}g'(\alpha_{nj})-\alpha_{nk}g'(\alpha_{nk})], \quad (81)$$

$$\int_0^1 r\,drf(\alpha_{nk}r)g(\alpha_{nk}r)=\tfrac{1}{4}[(f'(\alpha_{nk}))^2+(g'(\alpha_{nk}))^2]-\frac{(n+\frac{3}{2})^2}{{\alpha_{nk}}^2}, \quad (82)$$

$$\int_0^1 r\,drf(\alpha_{nk}r)g_{n/m}(\alpha_{mj}r)=\frac{1}{{\alpha_{nk}}^4-{\alpha_{nj}}^4}\{2{\alpha_{nk}}^2\alpha_{mj}\,g'_{n/m}(\alpha_{mj})$$
$$-{\alpha_{nk}}^3g'(\alpha_{nk})g_{n/m}(\alpha_{mj})-\alpha_{nk}{\alpha_{mj}}^2f'(\alpha_{nk})f_{n/m}(\alpha_{mj})\}. \quad . \quad . \quad (83)$$

With the help of the recursion relations for Bessel function we obtain

$$\int_0^1 r\,drf(\alpha_{nk}r)r^{n+3/2}=\frac{2n+3}{{\alpha_{nk}}^2}-\frac{g'(\alpha_{nk})}{\alpha_{nk}}. \quad . \quad . \quad . \quad (84)$$

Equations (78) through (84) and further use of recursion relations enables us to write the rest of the matrix elements as

$$\langle nk|FT|nj\rangle=\frac{2(n+1)^2(n+2)^2}{2n+3}\left\{-\frac{2n+3}{{\alpha_{nk}}^2{\alpha_{nj}}^2}\right.$$
$$\left.+\frac{2}{{\alpha_{nk}}^4-{\alpha_{nj}}^4}\left[{\alpha_{nk}}^2\frac{g'(\alpha_{nj})}{\alpha_{nj}}-{\alpha_{nj}}^2\frac{g'(\alpha_{nk})}{\alpha_{nk}}\right]\right\}, \quad (85)$$

$$\langle nk|FT|nk\rangle=\frac{2(n+1)^2(n+2)^2}{(2n+3){\alpha_{nk}}^2}\left\{\tfrac{1}{4}[(f'(\alpha_{nk}))^2+(g'(\alpha_{nk}))^2]\right.$$
$$\left.-\frac{n^2+7n+33/4}{{\alpha_{nk}}^2}+\frac{2g'(\alpha_{nk})}{\alpha_{nk}}\right\}, \quad (86)$$

$$\langle nk|FQ|nj\rangle=p\frac{2(n+1)(n+2)(n+3)}{(2n+3)^2}\left(\frac{2n+3}{{\alpha_{nk}}^2}-\frac{g'(\alpha_{nk})}{\alpha_{nk}}\right)\left(\frac{2n+3}{{\alpha_{nj}}^2}-\frac{g'(\alpha_{nj})}{\alpha_{nj}}\right)$$
$$+\frac{2n^2+6n+1}{(2n+1)(2n+5)}\langle nk|F^2|nj\rangle, \quad (87)$$

$$\langle nk|FP|n-2,j\rangle=\frac{2(n-1)n(n+1)(n+2)}{(2n-1)(2n+3)}\frac{1}{{\alpha_{nk}}^4-{\alpha_{n-2,j}}^4}$$
$$\times\left\{\alpha_{nk}\alpha_{n-2,j}f'(\alpha_{nk})f'(\alpha_{n-2,j})\right.$$
$$\left.-{\alpha_{nk}}^4\left(\frac{2n+3}{{\alpha_{nk}}^2}+\frac{g'(\alpha_{nk})}{\alpha_{nk}}\right)\left(\frac{2n-1}{{\alpha_{n-2,j}}^2}-\frac{g'(\alpha_{n-2,j})}{\alpha_{n-2,j}}\right)\right\}, \quad (88)$$

$$\langle nk|FR|n+2,j\rangle = \frac{2(n+1)(n+2)(n+3)(n+4)}{(2n+3)(2n+7)} \frac{1}{{\alpha_{nk}}^4 - {\alpha_{n+2,j}}^4}$$
$$\times \left\{ -\alpha_{nk}\alpha_{n+2,j} f'(\alpha_{nk}) f'(\alpha_{n+2,j}) \right.$$
$$\left. + {\alpha_{n+2,j}}^4 \left(\frac{2n+7}{{\alpha_{n+2,j}}^2} + \frac{g'(\alpha_{n+2,j})}{\alpha_{n+2,j}} \right) \left(\frac{2n+3}{{\alpha_{nk}}^2} - \frac{g'(\alpha_{nk})}{\alpha_{nk}} \right) \right\}. \quad (89)$$

The evaluation of the coefficients in equ. (73) for approximations up to nine by nine is given in the Appendix. In these tables a factor of two has been cancelled.

§ 8. Numerical Results

Because of the rather tedious calculations involved, it was decided to approximate equ. (73) with determinants of order less than or equal to nine. As may be seen by examining the apparent convergence of the values of **R** in table 1, eight parameters give a sufficient approximation for values of **T** less than 10^6 in the rigid boundary case and less than 10^5 for the free boundary.

The values of n and k assumed in our approximations are: $n=0$, 2 and 4 and $k=1$, 2 and 3.

Table 1. The Value of **R** for the Onset of the Lowest Mode of Instability for Various Values of the Taylor Number

For the case of a rigid boundary

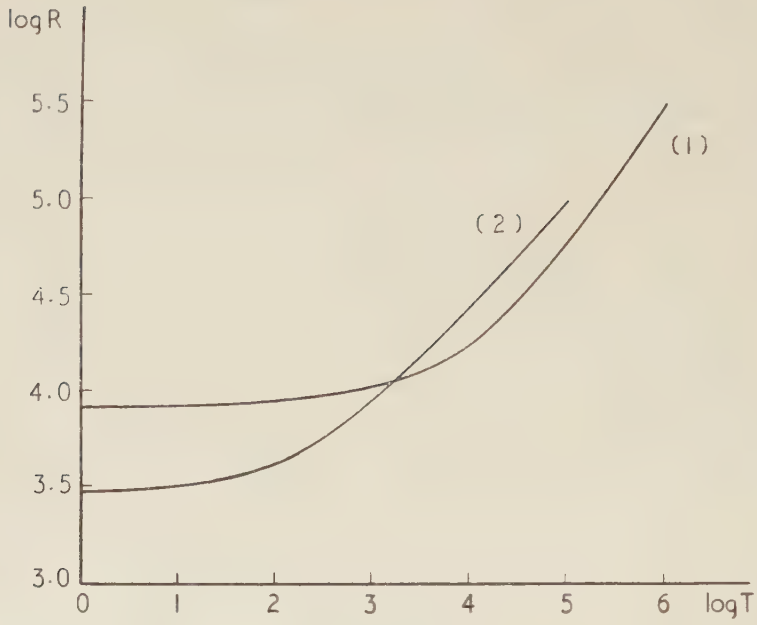
T	1st approx.	2nd approx.	4th approx.	6th approx.	9th approx.
0	$8{\cdot}0611\times10^3$	$8{\cdot}0421\times10^3$	—	—	—
10^3	$9{\cdot}4882\times10^3$	$9{\cdot}3855\times10^3$	$9{\cdot}3383\times10^3$	—	—
10^4	—	$1{\cdot}7306\times10^4$	$1{\cdot}6991\times10^4$	—	—
10^5	—	—	$0{\cdot}56993\times10^5$	$0{\cdot}53344\times10^5$	—
10^6	—	—	—	$0{\cdot}32122\times10^6$	$0{\cdot}29105\times10^6$

For the case of a free boundary

T	1st approx.	2nd approx.	4th approx.	6th approx.	9th approx.
0	$3{\cdot}0916\times10^3$	$3{\cdot}0912\times10^3$	—	—	—
10^3	$9{\cdot}4715\times10^3$	$9{\cdot}2493\times10^3$	$8{\cdot}9870\times10^3$	—	—
10^4	—	—	$2{\cdot}6395\times10^4$	$2{\cdot}5517\times10^4$	—
10^5	—	—	—	$0{\cdot}96961\times10^5$	$0{\cdot}91229\times10^5$

The results of the above tables are shown graphically in fig. 1.

Fig. 1



The dependence of the critical Rayleigh number for the onset of convective instability on the Taylor number in a rotating fluid sphere. The curves labelled (1) and (2) are for the cases of a rigid and a free bounding surface.

Table 2. The Characteristic Vectors Describing the Lowest Mode of Instability for Various Values of the Taylor number

For the case of a rigid boundary

	T = 0	$=10^3$	$=10^4$	$=10^5$	$=10^6$
A_{01}	0·99982	0·99859	0·97204	0·82340	0·85096
A_{02}	0·01917	0·02847	0·00342	−0·33881	0·37755
A_{03}	—	—	—	—	0·02874
A_{21}	—	0·04470	0·23249	0·42816	0·33137
A_{22}	—	0·00362	—	0·03278	−0·00423
A_{23}	—	—	—	—	0·00159
A_{41}	—	—	0·03268	0·14274	0·13141
A_{42}	—	—	—	0·04958	0·05937
A_{43}	—	—	—	—	0·03938

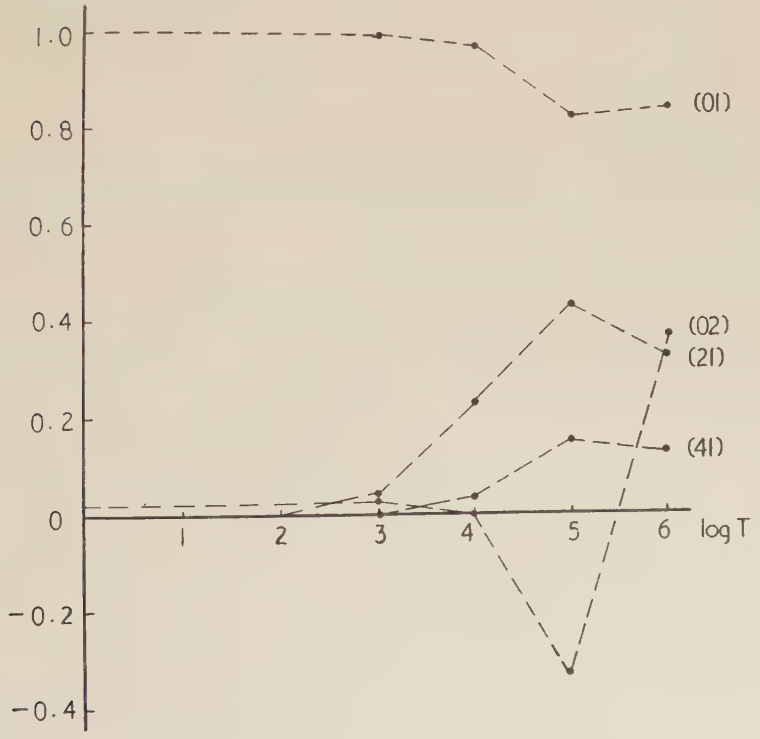
Table 2 (*cont.*)

For the case of a free boundary

	T = 0	$=10^3$	$=10^4$	$=10^5$
A_{01}	1·00000	0·99662	0·80118	0·61020
A_{02}	−0·00151	−0·04190	−0·36978	0·63219
A_{03}	—	−0·00388	−0·01094	−0·11656
A_{21}	—	0·07050	0·45896	−0·39936
A_{22}	—	—	−0·03995	0·16196
A_{23}	—	—	—	0·01481
A_{41}	—	—	0·09497	−0·16842
A_{42}	—	—	—	0·00975
A_{43}	—	—	—	0·00385

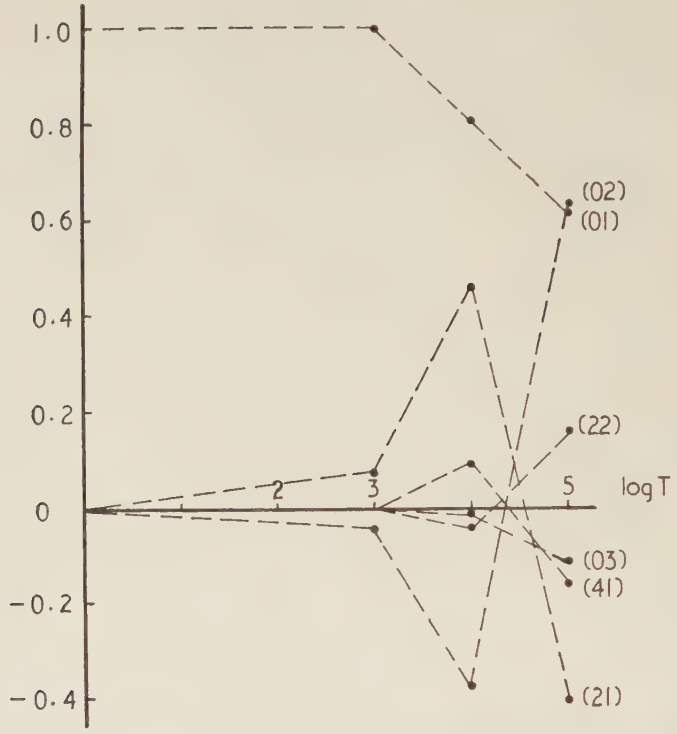
Though these are too few values for any quantitative conclusions, we have graphed the results in figs. 2 and 3 to show qualitatively the behaviour of a few of the components.

Fig. 2



The dependence of the components of the characteristic vector describing the lowest mode of instability on the Taylor number in a rotating fluid sphere with a rigid boundary.

Fig. 3



The dependence of the components of the characteristic vector describing the lowest mode of instability on the Taylor number in a rotating fluid sphere with a free boundary.

§ 9. Concluding Remarks

Because of the rather slow convergence of the successive approximations to the Rayleigh number for higher Taylor numbers, it is felt that further computations using the method described herein would not be worthwhile for this problem. Since the only self-gravitating spheres we can observe are of astronomical dimensions and (because of the dependence of **T** on the fourth power of the radius) have rather large Taylor numbers (e.g. $\sim 10^{26}$ for the earth's core); it is evident that an asymptotic relation giving the dependence of **R** on **T** when **T** is very large is needed.

A second point which may be mentioned is that there is certainly no good reason for assuming that instability does not appear as overstability rather than convection. Indeed, it has been shown that in the case of the plane Benard problem with rotation, instability *does* generally set in as overstability in the case of liquid metals such as mercury (Chandrasekhar and Elbert 1955, Fultz and Nakagawa 1955). Before any more extensive calculations are attempted for the problem in a sphere, the overstable case should be thoroughly investigated.

Acknowledgments

I would like finally to express my gratitude to Dr. Chandrasekhar for the aid and advice he has extended and to Mr. S. K. Trehan for reading the manuscript and checking the formulae.

The research reported in this paper has in part been supported by the Geophysics Research Directorate of the Air Force Cambridge Research Centre, Air Research and Development Command, under Contract AF 19 (604)–2046 with the University of Chicago.

References

Chandrasekhar, S., 1952, *Phil. Mag.*, **43**, 1317; 1956 a, *Proc. roy. Soc.* A, **237**, 476; 1956 b, *Astrophys. J.*, **124**, 232; 1957 a, *Phil. Mag.*, **2**, 845; 1957 b, *Ibid.*, **2**, 1282.

Chandrasekhar, S., and Elbert, D. D., 1955, *Proc. roy. Soc.* A, **231**, 198.

Chandrasekhar, S., and Reid, W. H., 1957, *Proc. nat. Acad. Sci., Wash.*, **43**, 521.

Erdelyi, A., Magnus, W., Oberhettinger, F., and Tricomi, F. G., **1953**, *Higher Transcendental Functions*, Vol. 2 (New York: McGraw-Hill).

Fultz, D., and Nakagawa, Y., 1955, *Proc. roy. Soc.* A, **231**, 211.

Lüst, R., and Schlüter, A., 1954, *Zs. f. Ap.*, **34**, 263.

Sommerfeld, A., 1949, *Partial Differential Equations* (New York: Academic Press Inc.).

APPENDIX

Further Tables of Numerical Results

Table 3 (*a*). Roots of the Equation $f(\alpha_{nj}) = 0$

j	$n=0$	$n=2$	$n=4$
1	5·267568	7·748590	10·10830
2	8·506951	11·19090	13·73380
3	11·68768	14·47661	17·12845

Table 3 (*b*). Roots of the Equation $f(\alpha_{nj}) = -\alpha_{nj}$

j	$n=0$	$n=2$	$n=4$
1	4·180881	6·813002	9·232055
2	7·573795	10·309365	12·881869
3	10·802354	13·618716	16·289137

Table 4. Numerical Values of $\langle nk|F^2|nk\rangle$

n, k	Rigid boundary	Free boundary
01	1·333333	6·330580
02	1·333333	28·60704
03	1·333333	63·92143
21	3·428571	54·65893
22	3·428571	146·4251
23	3·428571	271·3658
41	5·454545	175·8249
42	5·454545	378·4838
43	5·454545	632·3310

Table 5. The Numerical Value of the Constant Appearing on the Principal Diagonal in the Basic Determinants

n, k	Rigid boundary	Free boundary
01	$5{\cdot}13309\times10^{2}$	$9{\cdot}67133\times10^{2}$
02	$3{\cdot}49144\times10^{3}$	$4{\cdot}70649\times10^{4}$
03	$1{\cdot}24401\times10^{4}$	$4{\cdot}35201\times10^{5}$
21	$6{\cdot}17979\times10^{3}$	$5{\cdot}88825\times10^{4}$
22	$2{\cdot}68870\times10^{4}$	$8{\cdot}27014\times10^{5}$
23	$7{\cdot}52923\times10^{4}$	$4{\cdot}66735\times10^{6}$
41	$2{\cdot}84735\times10^{4}$	$6{\cdot}38621\times10^{5}$
42	$9{\cdot}70268\times10^{4}$	$5{\cdot}21114\times10^{6}$
43	$2{\cdot}34748\times10^{5}$	$2{\cdot}22590\times10^{7}$

The constant term appearing *off* the principal diagonal is identically zero.

Table 6 (*a*). The Coefficient of **R** in the Basic Determinant for a Rigid Boundary

n, k \ *m, j*	01	02	03	21	22	23	41	42	43
01	−0·063677	−0·0078633	−0·0045465	0	0	0	0	0	0
02	−0·0078633	−0·023059	−0·0029212	0	0	0	0	0	0
03	−0·0045465	−0·0029212	−0·011727	0	0	0	0	0	0
21	0	0	0	−0·40795	−0·039813	−0·026292	0	0	0
22	0	0	0	−0·039813	−0·19109	−0·018857	0	0	0
23	0	0	0	−0·026292	−0·018857	−0·11204	0	0	0
41	0	0	0	0	0	0	−0·90897	−0·072668	−0·051739
42	0	0	0	0	0	0	−0·072668	−0·49064	−0·039191
43	0	0	0	0	0	0	−0·051739	−0·039191	−0·30922

Table 6 (*b*). The Coefficient of **R** in the Basic Determinant for a Free Boundary

n, k \ *m, j*	01	02	03	21	22	23	41	42	43
01	−0·31283	0·024137	0·013744	0	0	0	0	0	0
02	0·024137	−0·47593	0·011566	0	0	0	0	0	0
03	0·0013744	0·011566	−0·53880	0	0	0	0	0	0
21	0	0	0	−6·7101	0·21907	0·14546	0	0	0
22	0	0	0	0·21907	−8·1055	0·11851	0	0	0
23	0	0	0	0·14546	0·11851	−8·4069	0	0	0
41	0	0	0	0	0	0	−30·1170	0·56509	0·40672
42	0	0	0	0	0	0	0·56509	−33·7805	0·33442
43	0	0	0	0	0	0	0·40672	0·33442	−35·4731

Table 7 (*a*). The Coefficient of **T** in the Basic Determinant for a Rigid Boundary

n, k \ m, j	01	02	03	21	22	23	41	42	43
01	0·0908693	−0·030876	−0·024095	−0·141558	−0·071521	−0·050194	0	0	0
02	−0·030876	0·110883	−0·017520	0·178765	−0·074242	−0·040712	0	0	0
03	−0·024095	−0·017520	0·119661	0·015125	0·203384	−0·048653	0	0	0
21	−0·141558	0·178765	0·015125	0·714257	−0·069591	−0·058462	−0·355002	−0·167390	−0·117059
22	−0·071521	−0·074242	0·203384	−0·069591	0·743517	−0·047450	0·375581	−0·218357	−0·111313
23	−0·050194	−0·040712	−0·048653	−0·058462	−0·047450	0·760140	0·049417	0·437310	−0·156147
41	0	0	0	−0·355002	0·375581	0·049417	1·22351	−0·089757	−0·078360
42	0	0	0	−0·167390	−0·218357	0·437310	−0·089757	1·25206	−0·066884
43	0	0	0	−0·117059	−0·111313	−0·156147	−0·078360	−0·066884	1·26580

Table 7 (*b*). The Coefficient of **T** in the Basic Determinant for a Free Boundary

n, k \ m, j	01	02	03	21	22	23	41	42	43
01	1·99585	1·52575	1·61540	−1·01187	−0·81272	−0·78112	0	0	0
02	1·52575	4·56891	1·80857	3·60901	−1·82257	−1·21594	0	0	0
03	1·61540	1·80857	8·30700	0·63765	9·58563	−2·49781	0	0	0
21	−1·01187	3·60901	0·63765	17·9097	5·35157	5·48525	−7·11288	−4·69060	−4·20665
22	−0·81272	−1·82257	9·58563	5·35157	39·7205	5·69345	15·0996	−10·3847	−6·26109
23	−0·78112	−1·21594	−2·49781	5·48525	5·69345	69·1543	3·42350	33·8452	−13·1723
41	0	0	0	−7·11288	15·0996	3·42350	53·7153	11·0965	11·2524
42	0	0	0	−4·69060	−10·3847	33·8452	11·0965	103·5057	11·4700
43	0	0	0	−4·20665	−6·26109	−13·1723	11·2524	11·4700	165·6605

Electrons in Polar Crystals†

By G. L. Sewell
Department of Applied Mathematics, The University, Liverpool

[Received July 16, 1958]

Abstract

The theory of the properties of an electron in a polar crystal is investigated by means of a model which takes account of the atomicity of the lattice. The method used is essentially that of a tight-binding approximation in which the interaction of the electron with the lattice displacements is built into the theory from the start. This method is valid only in cases where the tight-binding approach is appropriate even when the effects due to the lattice displacements are ignored.

Thus, the eigenstates of the system are constructed as linear combinations of localized states. In each of the localized states, the electron is bound to one positive ion, and displaces the neighbouring ions. The transfer of the electron between ions is therefore accompanied by a transfer of the mean position of the neighbouring ions. The effective mass of the electron is consequently increased as a result of the displacements it carries with itself. It is found that this effective mass increases exponentially with the electron–phonon coupling and can thus be extremely large in the strong coupling case. It is also found that the effective mass increases with temperature, since the random thermal motion of the ions opposes the motion of the lattice displacements accompanying the electron.

§ 1. Introduction and Discussion

An electron in a static periodic lattice behaves, in many respects, like a free particle, whose effective mass, m^0, depends on the potential due to the lattice. In polar crystals, the interaction of an electron with the polarization field, generated by the lattice vibrations, increases its effective mass to m^*, the polaron effective mass.

The existing theories of the polaron are based on a model due to Fröhlich (Fröhlich 1954, Allcock 1956), in which the polarization field is treated as a continuum, and in which the electron is treated as a free particle, of mass m^0. According to these theories, the polaron effective mass is not much larger than m^0 unless m^0 is itself large. In this latter case, however, the continuum model breaks down, since it leads to a picture in which the electron behaves like a wave-packet that is localized within a distance less than the lattice spacing.

A different approach to the theory of polarons of very high effective mass has been discussed by Fröhlich (1957), in connection with the Debye type of dielectric losses that occur in some ionic crystals, due to electrons

† Communicated by Professor H. Fröhlich, F.R.S.

bound to lattice defect centres (e.g. colour centres). Fröhlich considers cases in which the original effective mass m^0 is large. These are precisely the cases where the Bloch tight-binding approximation would be applicable if the lattice displacements were neglected. The essence of the new approach is that, in such cases, one may also calculate the eigenstates of the electron-lattice system by means of a tight-binding approximation in which the interaction of the electron with the lattice displacements is built into the theory from the start.

Thus, in determining the properties of an electron bound to a defect centre, one starts from a state in which the electron is localized on one of the positive ions near the centre and displaces the positions of the remaining ions. There will be a number of equivalently placed positive ions near the centre, and hence there will be a degenerate set of localized states of the above type. The energy required to localize the electron on any one of these ions is simply the overlap integral governing the transfer of the electron, together with its accompanying displacements, between neighbouring positive ions. This integral is given by

$$W \simeq W^0 \exp(-A^2/A_0^2), \quad . \quad . \quad . \quad . \quad . \quad . \quad (1.1)$$

where W^0 is the value the overlap integral would have if the effect of the lattice displacements were ignored, and the exponential factor arises from the transfer of the lattice displacements. In this latter factor, A is the displacement in the mean position of a positive ion, due to the presence of an electron on a neighbouring ion, and A_0 is the amplitude of the optical zero point vibrations of the positive ions. It is clear that the exponential factor may become extremely small in the case of strong coupling. In this case, the energy W, required to localize an electron on a positive ion near the defect, may be less than κT. If this is so, the electron will behave like a *classical* charged particle which is confined to the positive ions in the neighbourhood of the defect. On application of an electric field, the electron will therefore behave like a classical dipole, and will consequently lead to Debye dielectric losses.

In the present paper, we shall employ the same tight-binding approach to determine the theory of polarons of very high effective mass, for the case when there are no lattice defects. Our theory should therefore be applicable, for example, to the experiments of Morin (1954), which have shown that the electronic mobility in some ionic crystals (e.g. NiO and α-Fe_2O_3) are extraordinarily low. Such low mobilities can presumably only be explained on the basis of an extremely high effective mass.

Our theory will be based on a model in which an electron interacts with a polar lattice whose basic structure is that of two interlocking cubic sub-lattices (one for each set of ions). In order to determine the eigenstates of the system, we first obtain a localized state in which the electron is bound to one positive ion, whose motion it follows adiabatically. In this state, the presence of the electron on one ion will lead to displacements in the positions of the neighbouring ions. Moreover, there will be

a degenerate set of such states, one corresponding to each positive ion. By taking linear combinations of these localized states, we may remove the degeneracy and obtain the eigenstates of the system. The energy levels corresponding to these eigenstates will form a band, whose width is given by the overlap integral governing the transfer of the electron, and its accompanying displacements, between neighbouring positive ions. This band-width will therefore be given by (1.1), provided that we now interpret W as the polaron band-width and W^0 as the corresponding electronic band-width when the lattice displacements are ignored. We see from this formula that the polaron band may be very narrow, and thus the polaron effective mass may be very large, if $A \gg A_0$. As we shall show in § 7, this criterion can be satisfied for reasonable values of the characteristic constants of the model.

The polaron band-width is temperature dependent, and is given by the thermal average of the overlap integrals governing the transfer of the electron and its accompanying displacements. In fact it decreases with temperature, which means that the polaron effective mass increases with temperature. The essential reason for this is that the random thermal motion of the ions opposes the transfer of their mean positions.

It should be pointed out that the electrons whose low mobilities were measured experimentally were apparently d-electrons. However, in order to simplify the mathematics, we have confined our theory to a polaron band based on electronic s-states. The present theory can therefore be regarded only as a qualitative, rather than a quantitative, approach to the problem of the properties of electrons in polar crystals of the type mentioned above.

In a second paper, it will be shown that our theory leads to the prediction of a very low electronic mobility.

§ 2. The Model

We shall employ a model consisting of an electron, of mass m_{el}, moving through a crystal built up of positive ions, of mass M, and negative ions, of mass M'. It is convenient to introduce the total mass and the reduced mass, for a pair of opposite ions, i.e.

$$M_a = M + M' \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.1)$$

and

$$M_0 = \frac{MM'}{M+M'}\,; \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.2)$$

and also the reduced mass for an electron moving in the field of a positive ion, i.e.

$$m = \frac{Mm_{\mathrm{el}}}{M+m_{\mathrm{el}}}\,. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.3)$$

The mean positions of the positive and negative ions will be denoted by $\mathbf{a}_s$, $\mathbf{a}_s'$ respectively, where

$$\mathbf{a}_s \equiv a\sum_{l=1}^{3} s_l \mathbf{u}_l \equiv a(s_1, s_2, s_3), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.4)$$

the numbers s_1, s_2, s_3 being integers, and $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ forming a mutually orthogonal set of unit vectors; and

$$\mathbf{a}_s' = \mathbf{a}_s + a\boldsymbol{\alpha}, \quad \boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3), \quad . \quad . \quad . \quad . \quad . \quad (2.5)$$

the three components of $\boldsymbol{\alpha}$ all lying between 0 and 1, and being independent of s.

The position of the electron will be denoted by $\mathbf{x}$, and the displacements, from their mean positions, of the sth positive and negative ions will be denoted by $\mathbf{X}_s$, $\mathbf{X}_s'$, respectively. The relative displacement of the sth ions, and the mass centre corresponding to the displacements of these ions are thus

$$\mathbf{y}_s = \mathbf{X}_s - \mathbf{X}_s' \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.6)$$

and

$$\mathbf{z}_s = \frac{M\mathbf{X}_s + M'\mathbf{X}_s'}{M+M'}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.7)$$

It follows from these equations that we may express $\mathbf{X}_s$, $\mathbf{X}_s'$ in terms of $\mathbf{y}_s$, $\mathbf{z}_s$ by the equations

$$\mathbf{X}_s = \mathbf{z}_s + g\mathbf{y}_s \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.8)$$

and

$$\mathbf{X}_s' = \mathbf{z}_s - (1-g)\mathbf{y}_s, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.9)$$

where

$$g = \frac{M'}{M+M'}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.10)$$

It will be assumed that the optical and acoustical modes are described by the $\mathbf{y}$'s and $\mathbf{z}$'s, respectively, though this is really the case only for the long waves (cf. Born and Huang 1954, p. 82).

The Hamiltonian for the system is

$$H = H(\mathrm{el}) + H(\mathrm{lat}) + H(\mathrm{int}), \quad . \quad . \quad . \quad . \quad . \quad (2.11)$$

the three parts referring to the electron, the lattice and the electron–lattice interaction. Here,

$$H(\mathrm{el}) = -\frac{\hbar^2}{2m_{\mathrm{el}}}\frac{\partial^2}{\partial \mathbf{x}^2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.12)$$

The interaction $H(\mathrm{int})$ is the potential energy of the electron in the field due to the ions. Thus.

$$H(\mathrm{int}) = \sum_s [V(\mathbf{x} - \mathbf{a}_s - \mathbf{X}_s) + V'(\mathbf{x} - \mathbf{a}_s' - \mathbf{X}_s')], \quad . \quad . \quad (2.13)$$

where $V(\mathbf{x}-\mathbf{X})$, $V'(\mathbf{x}-\mathbf{X}')$ denote respectively the potentials at $\mathbf{x}$ due to a positive ion at $\mathbf{X}$ and to a negative ion at $\mathbf{X}'$.

We shall assume that the lattice Hamiltonian may be split up into the form

$$H(\mathrm{lat}) = H_{\mathrm{opt}} + H_{\mathrm{ac}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.14)$$

the two contributions to $H(\text{lat})$ being due to the optical and acoustical vibrations, respectively. These contributions may be written as

$$H_{\text{opt}}=\sum_s -\frac{\hbar^2}{2M_0}\frac{\partial^2}{\partial \mathbf{y}_s{}^2}+V_{\text{opt}}(\mathbf{y}), \quad . \quad . \quad . \quad . \quad (2.15)$$

and

$$H_{\text{ac}}=\sum_s -\frac{\hbar^2}{2M_a}\frac{\partial^2}{\partial \mathbf{z}_s{}^2}+V_{\text{ac}}(\mathbf{z}), \quad . \quad . \quad . \quad . \quad (2.16)$$

where V_{opt} and V_{ac} are functions of the $\mathbf{y}_s$'s and $\mathbf{z}_s$'s respectively. The formulae (2.14)–(2.16) rest on the hypothesis that the total potential energy of the lattice can be written in the form $V_{\text{opt}}(\mathbf{y})+V_{\text{ac}}(\mathbf{z})$. It is easily seen that the lattice kinetic energy, as given by (2.14)–(2.16), is

$$\sum_s\left[-\frac{\hbar^2}{2M_0}\frac{\partial^2}{\partial \mathbf{y}_s{}^2}-\frac{\hbar^2}{2M_a}\frac{\partial^2}{\partial \mathbf{z}_s{}^2}\right],$$

which, in view of (2.1), (2.2), (2.6) and (2.7), is equal to

$$\sum_s\left[-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial \mathbf{X}_s{}^2}-\frac{\hbar^2}{2M'}\frac{\partial^2}{\partial \mathbf{X}_s{}'^2}\right],$$

the correct formula for the total kinetic energy of the ions.

The Hamiltonians H_{opt} and H_{ac} may now be simplified by writing the $\mathbf{y}$'s and $\mathbf{z}$'s in terms of normal coordinates. We shall neglect the transverse modes, since these interact less strongly with the electron than do the longitudinal modes. Thus (cf. Wilson 1953, p. 251),

$$\mathbf{y}_s=\left(\frac{2}{NM_0}\right)^{1/2}\sum_{\mathbf{w}}{}'[\xi_{\mathbf{w}1}\cos \mathbf{w}\,.\,\mathbf{a}_s+\xi_{\mathbf{w}2}\sin \mathbf{w}\,.\,\mathbf{a}_s]\,\hat{\mathbf{w}} \quad . \quad . \quad (2.17)$$

and

$$\mathbf{z}_s=\left(\frac{2}{NM_a}\right)^{1/2}\sum_{\mathbf{w}}{}'[\xi_{\mathbf{w}3}\cos \mathbf{w}\,.\,\mathbf{a}_s+\xi_{\mathbf{w}4}\sin \mathbf{w}\,.\,\mathbf{a}_s]\,\hat{\mathbf{w}}. \quad . \quad (2.18)$$

In these equations, the ξ's are normal coordinates, N is the total number of pairs of ions, and the primes over the $\sum$'s denote summation over half of the occupied region in $\mathbf{w}$-space. This region is given by

$$\left.\begin{aligned} -\frac{\pi}{a}\leqslant w_1, w_2, w_3\leqslant \frac{\pi}{a} \\ (w_1, w_2, w_3)\equiv \mathbf{w}. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.19)$$

where

The Hamiltonians H_{opt} and H_{ac} may now be expressed in the forms

$$\left.\begin{aligned} H_{\text{opt}}&=\sum_{j=1}^{2}\sum_{\mathbf{w}}{}'\tfrac{1}{2}[\eta_{\mathbf{w}j}{}^2+\nu_{\mathbf{w}j}{}^2\xi_{\mathbf{w}j}{}^2] \\ \text{and}\quad H_{\text{ac}}&=\sum_{j=3}^{4}\sum_{\mathbf{w}}{}'\tfrac{1}{2}[\eta_{\mathbf{w}j}{}^2+\nu_{\mathbf{w}j}{}^2\xi_{\mathbf{w}j}{}^2], \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (2.20)$$

where $\eta_{\mathbf{w}j}$ is the momentum conjugate to $\xi_{\mathbf{w}j}$, and $\nu_{\mathbf{w}j}$ is the frequency for the mode $\mathbf{w}, j$. Hence, we may rewrite (2.14) as

$$H(\text{lat})=\sum_{j=1}^{4}\sum_{\mathbf{w}}{}'H_{\mathbf{w}j}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.21)$$

where

$$H_{\mathbf{w}j}=\tfrac{1}{2}[\eta_{\mathbf{w}j}{}^2+\nu_{\mathbf{w}j}{}^2\xi_{\mathbf{w}j}{}^2], \quad . \quad . \quad . \quad . \quad . \quad (2.22)$$

and

$$\eta_{\mathbf{w}j}=-i\hbar\partial/\partial\xi_{\mathbf{w}j}. \quad . \quad . \quad . \quad . \quad . \quad (2.23)$$

We shall assume that

$$\left.\begin{aligned}&\nu_{\mathbf{w}j}=\nu, \text{ for } j=1, 2 \text{ (optical modes)}\\ \text{and}\quad &\nu_{\mathbf{w}j}=cw, \text{ for } j=3, 4 \text{ (acoustical modes),}\end{aligned}\right\} \quad . \quad . \quad (2.24)$$

where ν (the infra-red absorption frequency) and c (the speed of sound) are constants. These equations for the frequencies are certainly valid for the long waves.

The eigenstates and energy levels of $H(\text{lat})$ are given by

$$\Theta_n=\prod_{j=1}^{4}\prod_{\mathbf{w}}{}'\theta(n_{\mathbf{w}j}, \xi_{\mathbf{w}j}) \quad . \quad . \quad . \quad . \quad . \quad . \quad (2.25)$$

and

$$E_n(\text{lat})=\sum_{j=1}^{4}\sum_{\mathbf{w}}{}'(n_{\mathbf{w}j}+\tfrac{1}{2})\hbar\nu_{\mathbf{w}j}, \quad . \quad . \quad . \quad . \quad (2.26)$$

where $\theta(n_{\mathbf{w}j}, \xi_{\mathbf{w}j})$ is the eigenstate of $H_{\mathbf{w}j}$ in which there are $n_{\mathbf{w}j}$ quanta.

§ 3. Localized States

We shall now determine the wave functions for states in which the electron is bound to one positive ion, and displaces the positions of the remaining ions. (We shall ignore the polarization induced in the ions themselves.) For this purpose we split H into two parts,

$$H=H_s+U_s, \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.1)$$

where

$$U_s=\sum_{t\neq s}V(\mathbf{x}-\mathbf{a}_t)+\sum_t V'(\mathbf{x}-\mathbf{a}_t') \quad . \quad . \quad . \quad . \quad . \quad (3.2)$$

is the potential energy of the electron due to all the ions except the sth positive one, when they are in their mean positions. Thus, by (2.11), (2.13), (3.1) and (3.2),

$$\begin{aligned}H_s=H(\text{el})+V(\mathbf{x}-\mathbf{a}_s-\mathbf{X}_s)+H(\text{lat})+\sum_{t\neq s}[V(\mathbf{x}-\mathbf{a}_t-\mathbf{X}_t)-V(\mathbf{x}-\mathbf{a}_t)]\\ +\sum_t[V'(\mathbf{x}-\mathbf{a}_t'-\mathbf{X}_t')-V'(\mathbf{x}-\mathbf{a}_t')], \quad . \quad . \quad . \quad (3.3)\end{aligned}$$

which means that H_s is the Hamiltonian for a system in which an electron interacts with the sth positive ion and with the displacements of the remaining ions.

The localized states we shall consider will be stationary states of H_s, given by the form

$$\psi_s=\phi(\mathbf{x}-\mathbf{a}_s-\mathbf{X}_s)\chi_s(\text{lat}), \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.4)$$

where χ_s depends on the lattice coordinates ξ only. ψ_s is thus a state in which the electron follows the sth positive ion adiabatically.

It follows from (2·1)–(2.3), (2.6), (2.7), (2.11)–(2.16), (3.3) and (3.4) that

$$H_s\psi_s=\chi_s\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial\mathbf{x}^2}+V(\mathbf{x}-\mathbf{a}_s-\mathbf{X}_s)\right]\phi(\mathbf{x}-\mathbf{a}_s-\mathbf{X}_s)-\frac{\hbar^2}{M}\frac{\partial\phi}{\partial\mathbf{x}}\cdot\frac{\partial\chi_s}{\partial\mathbf{X}_s}$$
$$+\phi\left[H(\mathrm{lat})+\sum_{t\neq s}[V(\mathbf{x}-\mathbf{a}_t-\mathbf{X}_t)-V(\mathbf{x}-\mathbf{a}_t)]+\sum_t[V'(\mathbf{x}-\mathbf{a}_t'-\mathbf{X}_t')\right.$$
$$\left.-V'(\mathbf{x}-\mathbf{a}_t')]\right]\chi_s. \quad . \quad . \quad . \quad (3.5)$$

We shall assume ϕ to be an eigenstate of the electron in the potential due to the sth positive ion, thereby assuming that the lattice spacing is large enough for this atomic state not to be appreciably affected by the displacements of the other ions. Thus, we shall assume that

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial\mathbf{x}^2}+V(\mathbf{x}-\mathbf{a}_s-\mathbf{X}_s)\right]\phi(\mathbf{x}-\mathbf{a}_s-\mathbf{X}_s)=\epsilon_0\phi(\mathbf{x}-\mathbf{a}_s-\mathbf{X}_s), \quad (3.6)$$

ϵ_0 being the energy corresponding to the atomic state ϕ. Further, we shall stipulate that ϕ is an s-state, since the non-degeneracy of such atomic states leads to mathematical simplifications. χ_s, on the other hand, will be obtained from the condition that the energy of the system H_s for the state ψ_s is stationary. ϕ and χ_s will be chosen so as to be normalized with respect to the electron and lattice coordinates respectively. Hence, it follows from (3.5) and (3.6) that, since

$$\int\phi^+\nabla\phi\,d^3\mathbf{x}=\int\phi\nabla\phi\,d^3\mathbf{x}=0 \quad (\phi \text{ real}),$$

then the energy of the system H_s, for the state ψ_s, is

$$(\psi_s^+|H_s|\psi_s)=\epsilon_0+(\chi_s^+|H_s(\mathrm{lat})|\chi_s), \quad . \quad . \quad . \quad (3.7)$$

where

$$H_s(\mathrm{lat})=H(\mathrm{lat})+\int|\phi(\mathbf{x}-\mathbf{a}_s-\mathbf{X}_s)|^2\left[\sum_{t\neq s}[V(\mathbf{x}-\mathbf{a}_t-\mathbf{X}_t)-V(\mathbf{x}-\mathbf{a}_t)]\right.$$
$$\left.+\sum_t[V'(\mathbf{x}-\mathbf{a}_t'-\mathbf{X}_t')-V'(\mathbf{x}-\mathbf{a}_t')]\right]d^3\mathbf{x}. \quad . \quad . \quad . \quad (3.8)$$

It follows from (3.7) that, if ψ_s is a stationary state of H_s, then χ_s is an eigenstate of $H_s(\mathrm{lat})$, with energy, say E_s' so that

$$H_s(\mathrm{lat})\chi_s=E_s'\chi_s \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.9)$$

and

$$(\psi_s^+|H_s|\psi_s)=E_s'+\epsilon_0. \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.10)$$

We may simplify the formula (3.8) for $H_s(\mathrm{lat})$ by neglecting all powers of the $\mathbf{X}$'s and $\mathbf{X}'$'s higher than the first in the integral appearing in that formula. Hence,

$$H_s(\mathrm{lat})=H(\mathrm{lat})-\sum_{t\neq s}\mathbf{X}_t\cdot\int|\phi(\mathbf{x}-\mathbf{a}_s)|^2\nabla V(\mathbf{x}-\mathbf{a}_t)\,d^3\mathbf{x}$$
$$-\sum_t\mathbf{X}_t'\cdot\int|\phi(\mathbf{x}-\mathbf{a}_s)|^2\nabla V'(\mathbf{x}-\mathbf{a}_t')\,d^3\mathbf{x}.$$

This formula can be rewritten, in view of the lattice periodicity as

$$H_s(\mathrm{lat})=H(\mathrm{lat})-\sum_t\left[\mathbf{X}_t\cdot\int|\phi(\mathbf{x})|^2\nabla V(\mathbf{x}-\mathbf{a}_t+\mathbf{a}_s)\,d^3\mathbf{x}+\mathbf{X}_t'\cdot\int|\phi(\mathbf{x})|^2\right.$$
$$\left.\times\nabla V'(\mathbf{x}-\mathbf{a}_t'+\mathbf{a}_s)\,d^3\mathbf{x}\right]+\mathbf{X}_s\cdot\int|\phi(\mathbf{x})|^2\nabla V(\mathbf{x})\,d^3\mathbf{x}.$$

The last term in this equation is zero, since $|\phi(\mathbf{x})|^2$ is an even function of $\mathbf{x}$, while $\nabla V(\mathbf{x})$ is an odd function of $\mathbf{x}$, since $V(\mathbf{x})$ is a radial function. Hence, putting

$$\left.\begin{aligned}\mathbf{B}(\mathbf{X}) &\equiv \int |\phi(\mathbf{x})|^2 \nabla[gv(\mathbf{x}-\mathbf{X}) - (1-g)V'(\mathbf{x}-\mathbf{X}+\mathbf{a}\alpha)]\, d^3\mathbf{x} \\ \text{and} \qquad \mathbf{C}(\mathbf{X}) &\equiv \int |\phi(\mathbf{x})|^2 \nabla[V(\mathbf{x}-\mathbf{X}) + V'(\mathbf{x}-\mathbf{X}+a\alpha)]\, d^3\mathbf{x},\end{aligned}\right\} \tag{3.11}$$

it follows from (2.5), (2.8) and (2.9) that

$$H_s(\text{lat}) = H(\text{lat}) - \sum_t [\mathbf{y}_t \cdot \mathbf{B}(\mathbf{a}_t - \mathbf{a}_s) + \mathbf{z}_t \cdot \mathbf{C}(\mathbf{a}_t - \mathbf{a}_s)].$$

The last equation may be rewritten, using (2.17), (2.18), (2.21), (2.22) and (2.24), as

$$H_s(\text{lat}) = \sum_{j=1}^{4} {\sum_{\mathbf{w}}}' \tfrac{1}{2}[\eta_{\mathbf{w}j}^2 + \nu_{\mathbf{w}j}^2(\xi_{\mathbf{w}j} - \bar{\xi}_{\mathbf{w}js})^2] - \epsilon_1, \tag{3.12}$$

where

$$\epsilon_1 = \sum_{j=1}^{4} {\sum_{\mathbf{w}}}' \tfrac{1}{2}\nu_{\mathbf{w}j}^2 \bar{\xi}_{\mathbf{w}js}^2, \tag{3.13}$$

where the $\bar{\xi}$'s are all real, and where

$$\bar{\xi}_{\mathbf{w}1s} + i\bar{\xi}_{\mathbf{w}2s} = \nu^{-2}\left(\frac{2}{NM_0}\right)^{1/2} \sum_t \hat{\mathbf{w}} \cdot \mathbf{B}(\mathbf{a}_t - \mathbf{a}_s) \exp(i\mathbf{w} \cdot \mathbf{a}_t)$$

and

$$\bar{\xi}_{\mathbf{w}3s} + i\bar{\xi}_{\mathbf{w}4s} = (cw)^{-2}\left(\frac{2}{NM_a}\right)^{1/2} \sum_t \hat{\mathbf{w}} \cdot \mathbf{C}(\mathbf{a}_t - \mathbf{a}_s) \exp(i\mathbf{w} \cdot \mathbf{a}_t),$$

i.e.

$$\left.\begin{aligned}\bar{\xi}_{\mathbf{w}1s} + i\bar{\xi}_{\mathbf{w}2s} &= N^{-1/2} B_{\mathbf{w}} \exp(i\mathbf{w} \cdot \mathbf{a}_s) \\ \text{and} \qquad \bar{\xi}_{\mathbf{w}3s} + i\bar{\xi}_{\mathbf{w}4s} &= N^{-1/2} C_{\mathbf{w}} \exp(i\mathbf{w} \cdot \mathbf{a}_s),\end{aligned}\right\} \tag{3.14}$$

where

$$\left.\begin{aligned}B_{\mathbf{w}} &= \nu^{-2}\left(\frac{2}{M_0}\right)^{1/2} \sum_t \hat{\mathbf{w}} \cdot \mathbf{B}(\mathbf{a}_t) \exp(i\mathbf{w} \cdot \mathbf{a}_t) \\ \text{and} \qquad C_{\mathbf{w}} &= (cw)^{-2}\left(\frac{2}{M_a}\right)^{1/2} \sum_t \hat{\mathbf{w}} \cdot \mathbf{C}(\mathbf{a}_t) \exp(i\mathbf{w} \cdot \mathbf{a}_t).\end{aligned}\right\} \tag{3.15}$$

By (2.21) and (2.23), we may rewrite (3.12) as

$$H_s(\text{lat}) = T_s H(\text{lat}) T_s^{+} - \epsilon_1, \tag{3.16}$$

where

$$T_s = \exp\left[-\frac{i}{\hbar} \sum_{j=1}^{4} {\sum_{\mathbf{w}}}' \bar{\xi}_{\mathbf{w}js}\eta_{\mathbf{w}j}\right] \tag{3.17}$$

is a unitary displacement operator. Consequently, the eigenstates and energy levels of $H_s(\text{lat})$ are related to those of $H(\text{lat})$ by the equations

$$\chi_s = \chi_{sn} = T_s \Theta_n \tag{3.18}$$

and

$$E_s' = E_n(\text{lat}) - \epsilon_1.$$

Hence, by (3.4), (3.9) and (3.10), the localized wave functions are given by

$$\psi_s = \psi_{sn} = \phi(\mathbf{x} - \mathbf{a}_s - \mathbf{X}_s)\chi_{sn}, \quad . \quad . \quad . \quad . \quad (3.19)$$

and the corresponding energy levels are

$$(\psi_{sn}^+ | H_s | \psi_{sn}) = E_n(\text{lat}) + \epsilon_0 - \epsilon_1. \quad . \quad . \quad . \quad . \quad (3.20)$$

§ 4. The Band Structure

For a given distribution of phonons (i.e. given Θ_n), there will be N states ψ_{sn}, one for each positive ion. These states form a degenerate set. By taking linear combinations of them we may obtain stationary states of the system. Thus, we shall seek stationary states of H, that of the form

$$\Psi = \sum_s c_s \psi_{sn}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.1)\dagger$$

By (3.1) and (4.1), the energy E of the system H, for the state Ψ, is given by

$$\sum_{st} c_s^+ c_t (\psi_{sn}^+ | E - H_s - U_s | \psi_{tn}) = 0.$$

Choosing the c's so that E is stationary, we obtain N secular equations

$$\sum_t c_t (\psi_{sn}^+ | E - H_s - U_s | \psi_{tn}) = 0. \quad . \quad . \quad . \quad . \quad . \quad (4.2)$$

In order to solve this set of equations, we adopt a tight-binding approximation (cf. Bloch 1928), and put

$$(\psi_{sn}^+ | \psi_{tn}) = \delta_{st}. \quad . \quad . \quad . \quad . \quad . \quad (4.3)$$

Thus, if we assume that the ψ_{sn}'s are eigenstates of H_s, it follows from (3.22) that

$$(\psi_{sn}^- | H_s | \psi_{tn}) = (E_n(\text{lat}) + \epsilon_0 - \epsilon_1)\delta_{st}. \quad . \quad . \quad . \quad . \quad (4.4)$$

The matrix elements $(\psi_{sn}^- | U_s | \psi_{tn})$ depend on $\mathbf{a}_s$, $\mathbf{a}_t$ only through their difference, in view of the lattice periodicity, so that we may put

$$(\psi_{sn}^+ | U_s | \psi_{tn}) = W_n(\mathbf{a}_t - \mathbf{a}_s). \quad . \quad . \quad . \quad . \quad . \quad (4.5)$$

By (4.3)–(4.5), we may rewrite (4.2) as

$$\sum_t c_t [(E + \epsilon_1 - \epsilon_0 - E_n(\text{lat}))\delta_{st} - W_n(\mathbf{a}_t - \mathbf{a}_s)] = 0.$$

The solution of this set of equations for E and the c's is

$$c_s = \exp(i\mathbf{k} \cdot \mathbf{a}_s), \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.6)$$

where

$$\mathbf{k} = (k_1, k_2, k_3). \quad . \quad . \quad . \quad . \quad . \quad . \quad (4.7)$$

is an electron wave vector which lies in the region

$$-\frac{\pi}{a} \leqslant k_1, k_2, k_3 \leqslant \frac{\pi}{a}; \quad . \quad . \quad . \quad . \quad . \quad (4.8)$$

† It is unnecessary to consider other terms in this sum, since the atomic energy level ϕ is non-degenerate, owing to the fact that ϕ has been chosen to be an s-state.

and the energy E is given by

$$E = \epsilon_0 - \epsilon_1 + E_n(\text{lat}) + \sum_t W_n(\mathbf{a}_t) \exp(i\,\mathbf{k} \cdot \mathbf{a}_t).$$

It follows from this last equation, together with (4.1) and (4.6), that the wave functions and energy levels under consideration are

$$\Psi = \Psi_{\mathbf{k}n} = \sum_s \exp(i\,\mathbf{k} \cdot \mathbf{a}_s)\psi_{sn} \quad . \quad . \quad . \quad . \quad (4.9)$$

and

$$E = E_{\mathbf{k}n} = E_n(\text{lat}) + \epsilon_{\mathbf{k}n}, \quad . \quad . \quad . \quad . \quad (4.10)$$

respectively, where

$$\epsilon_{\mathbf{k}n} = \epsilon_0 - \epsilon_1 + \sum_t W_n(\mathbf{a}_t) \exp(i\,\mathbf{k} \cdot \mathbf{a}_t). \quad . \quad . \quad . \quad (4.11)$$

Since $E_n(\text{lat})$ is the lattice energy, $\epsilon_{\mathbf{k}n}$ may be regarded as the polaron energy for the wave vector $\mathbf{k}$, when the lattice state is Θ_n.

The energies W_n may be expressed, using eqns. (3.2), (3.19) and (4.5), in the form

$$W_n(\mathbf{a}_t - \mathbf{a}_s) = (\chi_{sn}^+ \,|\, W_{st} \,|\, \chi_{tn}), \quad . \quad . \quad . \quad (4.12)$$

where

$$W_{st} = \int \phi^+(\mathbf{x} - \mathbf{a}_s - \mathbf{X}_s) U_s \phi(\mathbf{x} - \mathbf{a}_t - \mathbf{X}_t)\, d^3\mathbf{x}.$$

Assuming that the displacements $\mathbf{X}_s$, $\mathbf{X}_t$ are much less in magnitude than the lattice spacing, we may neglect these terms in the above formula for W_{st}. In this case,

$$W_{st} = \int \phi^+(\mathbf{x} - \mathbf{a}_s) U_s \phi(\mathbf{x} - \mathbf{a}_t)\, d^3\mathbf{x} = W^0(\mathbf{a}_t - \mathbf{a}_s), \quad . \quad . \quad (4.13)$$

a quantity depending on $\mathbf{a}_s$, $\mathbf{a}_t$ only through their difference. Consequently by (3.18) and (4.11)–(4.13),

$$\epsilon_{\mathbf{k}n} = (\Theta_n^+ \,|\, \epsilon_{\text{op}}(\mathbf{k}) \,|\, \Theta_n), \quad . \quad . \quad . \quad . \quad . \quad (4.14)$$

where the operator $\epsilon_{\text{op}}(\mathbf{k})$ is given by

$$\epsilon_{\text{op}}(\mathbf{k}) = \epsilon_0 - \epsilon_1 + \sum_t W^0(\mathbf{a}_t) T_0^+ T_t \exp(i\mathbf{k} \cdot \mathbf{a}_t). \quad . \quad . \quad (4.15)$$

§ 5. Temperature Dependence of Polaron Band

It follows from (4.14) that when the lattice state is Θ_n, the polaron energy is the expectation value of the operator $\epsilon_{\text{op}}(\mathbf{k})$ for that lattice state. Consequently, the polaron energy for the wave vector $\mathbf{k}$ at temperature T is the thermal average, taken over all lattice states, of $\epsilon_{\text{op}}(\mathbf{k})$. It is therefore equal to

$$\epsilon(\mathbf{k}) = \frac{Sp[\epsilon_{\text{op}}(\mathbf{k}) \exp(-\beta H(\text{lat}))]}{Sp[\exp(-\beta H(\text{lat}))]}, \quad . \quad . \quad . \quad . \quad (5.1)$$

where

$$\beta = (\kappa T)^{-1} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.2)$$

and κ is Boltzmann's constant. It will be shown in Appendix 3 that the mean square deviation from $\epsilon(\mathbf{k})$ of the energies $\epsilon_{\mathbf{k}n}$ is negligible for a macroscopic system. $\epsilon(\mathbf{k})$ may therefore be justifiably regarded as the polaron energy for the wave vector $\mathbf{k}$.

It follows from (4.15) and (5.1) that

$$\epsilon(\mathbf{k}) = \epsilon_0 - \epsilon_1 + \sum_t W(\mathbf{a}_t) \exp(i\mathbf{k} \cdot \mathbf{a}_t), \quad . \quad . \quad . \quad . \quad (5.3)$$

where

$$W(\mathbf{a}_t) = W^0(\mathbf{a}_t) F(\mathbf{a}_t) \quad . \quad . \quad . \quad . \quad (5.4)$$

and

$$F(\mathbf{a}_t) = \frac{Sp[T_0 + T_t \exp(-\beta H(\text{lat}))]}{Sp[\exp(-\beta H(\text{lat}))]} . \quad . \quad . \quad . \quad (5.5)$$

It will be shown in Appendix 2 that

$$W^+(\mathbf{a}_t) = W(\mathbf{a}_t) = W(-\mathbf{a}_t). \quad . \quad . \quad . \quad . \quad (5.6)$$

In view of this, eqn. (5.3) may be rewritten as

$$\epsilon(\mathbf{k}) = \epsilon_0 - \epsilon_1 + W(0) + 2\sum_t{}' W(\mathbf{a}_t) \cos \mathbf{k} \cdot \mathbf{a}_t,$$

the prime over the $\sum$ denoting summation over half of the lattice points $\mathbf{a}_t \neq 0$. We shall now ignore all the W's except those concerned with the transfer of the electron between nearest neighbouring positive ions. We shall therefore retain in our last equation for $\epsilon(\mathbf{k})$ only the terms for which

$$\mathbf{a}_t = a\mathbf{u}_1, a\mathbf{u}_2 \quad \text{or} \quad a\mathbf{u}_3,$$

the $\mathbf{u}$'s being defined in (2.4). Hence, using (4.7),

$$\epsilon(\mathbf{k}) = \epsilon_0 - \epsilon_1 + W(0) + 2\sum_{l=1}^{3} W_l \cos k_l a, \quad . \quad . \quad . \quad (5.7)$$

where

$$W_l = W_l^0 F_l, \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.8)$$

and

$$\left.\begin{aligned} F_l &= F(a\mathbf{u}_l) \\ W_l^0 &= W^0(a\mathbf{u}_l). \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.9)$$

It is easily seen from (5.7) that

$$\epsilon(\mathbf{k}) = \epsilon(0) - \sum_{l=1}^{3} W_l a^2 k_l^2 + 0(k^4).$$

Consequently, the effective mass of the polaron, for propagation in the direction of $\mathbf{u}_l$, is

$$m_l = -\frac{\hbar^2}{2W_l a^2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5.10)$$

§ 6. Calculation of m_l

It is seen from eqns. (4.13), (5.5), (5.8) and (5.9) that W_l is the product of two factors. The first of these, W_l^0, is the integral corresponding to the transfer of an electron between nearest neighbouring positive ions in a rigid lattice of the same basic structure as the one under consideration. It is related to the effective mass m_l^0, of an electron moving in the direction of $\mathbf{u}_l$, when the lattice is rigid, by the equation (cf. (5.10)),

$$m_l^0 = -\frac{\hbar^2}{2W_l^0 a^2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6.1)$$

The second factor, F_l, in the formula for W_l, arises as a result of the transfer from ion to ion of the lattice displacements that accompany the electron.

We may now express m_l in terms of F_l and m_l^0. For, by eqns. (5.10) and (6.1),

$$m_l = m_l^0/F_l. \qquad (6.2)$$

The effects of the lattice displacements on the polaron effective mass are therefore given by F_l. We shall now calculate this quantity.

By (2.21), (2.22), (3.17) and (5.5),

$$F(\mathbf{a}_l) = \prod_{j=1}^{\dagger} \prod_{\mathbf{w}}' f_{\mathbf{w}jl}, \qquad (6.3)$$

where

$$f_{\mathbf{w}jt} = \frac{Sp[\exp\{i(\bar{\xi}_{\mathbf{w}j0} - \bar{\xi}_{\mathbf{w}jt})\eta_{\mathbf{w}j}/\hbar\} \exp(-\beta H_{\mathbf{w}j})]}{Sp[\exp(-\beta H_{\mathbf{w}j})]}. \qquad (6.4)$$

It will be shown in Appendix 1 the solution of (6.4) is

$$f_{\mathbf{w}jt} = \exp\left[- \frac{\nu_{\mathbf{w}j}(\bar{\xi}_{\mathbf{w}j0} - \bar{\xi}_{\mathbf{w}jt})^2}{4\hbar} \coth(\tfrac{1}{2}\beta\hbar\nu_{\mathbf{w}j}) \right]. \qquad (6.5)$$

It now follows from (6.3) and (6.5) that

$$F(\mathbf{a}_l) = \exp\left[- \frac{1}{4\hbar} \sum_{j=1}^{4} \sum_{\mathbf{w}}' \nu_{\mathbf{w}j}(\bar{\xi}_{\mathbf{w}j0} - \bar{\xi}_{\mathbf{w}jl})^2 \coth(\tfrac{1}{2}\beta\hbar\nu_{\mathbf{w}j}) \right];$$

and therefore, by (2.24) and (3.14),

$$F(\mathbf{a}_l) = \exp\left[- \frac{1}{N\hbar} \sum_{j=1}^{4} \sum_{\mathbf{w}}' \sin^2 \tfrac{1}{2}\mathbf{w}.\mathbf{a}_l \{\nu |B_{\mathbf{w}}|^2 \coth(\tfrac{1}{2}\beta\hbar\nu) + cw|C_{\mathbf{w}}|^2 \coth(\tfrac{1}{2}\beta\hbar cw)\} \right].$$

Thus, by (5.2) and (5.9),

$$F_l = \exp(-\gamma_l) \qquad (6.6)$$

where

$$\gamma_l = \gamma_l(\text{opt}) + \gamma_l(\text{ac}), \qquad (6.7)$$

and where

$$\left.\begin{aligned} \gamma_l(\text{opt}) &= \frac{\nu}{N\hbar} \coth\left(\frac{\hbar\nu}{2\kappa T}\right) \sum_{\mathbf{w}}' |B_{\mathbf{w}}|^2 \sin^2 \tfrac{1}{2} a w_l \\ \text{and} \quad \gamma_l(\text{ac}) &= \frac{c}{N\hbar} \sum_{\mathbf{w}}' w |C_{\mathbf{w}}|^2 \sin^2 \tfrac{1}{2} a w_1 \coth\left(\frac{\hbar c w}{2\kappa T}\right). \end{aligned}\right\} \qquad (6.8)$$

$\gamma_l(\text{opt})$ and $\gamma_l(\text{ac})$ are thus dimensionless quantities which increase with temperature, and which depend on the coupling of the electron to the optical and acoustical vibrations respectively.

The polaron effective mass, m_l, and the energy W_l may now be expressed in terms of γ_l. For, by eqns. (5.8), (6.2) and (6.6),

$$m_l = m_l^0 \exp(\gamma_l) \qquad (6.9)$$

and

$$W_l = W_l^0 \exp(-\gamma_l). \qquad (6.10)$$

An order of magnitude approximation to this last equation, for the case $T=0$, in which the acoustical vibrations are ignored, is given by (1.1).

Equation (6.9) shows that the polaron effective mass increases exponentially with γ_l, and may thus be extremely large if

$$\gamma_l \gg 1. \qquad (6.11)$$

We shall show in the next section that this condition may be satisfied for reasonable values of the characteristic constants of the crystal.

Moreover, by (6.8) and (6.9), m_l increases with T; and it may be seen from those equations that

$$m_l = m_l^0 \exp(T/T_l), \quad \text{for } \kappa T \gg \hbar\nu, \hbar c/a, \qquad (6.12)$$

where

$$(\kappa T_l)^{-1} = \frac{2}{N\hbar^2} \sum_{\mathbf{w}}' \{|B_{\mathbf{w}}|^2 + |C_{\mathbf{w}}|^2\} \sin^2 \tfrac{1}{2} a w_l. \qquad (6.13)$$

The increase of m_l with T might have been anticipated from eqns. (5.9) and (6.2)–(6.4), according to which m_l is inversely proportional to the quantities $f_{\mathbf{w}jl}$, which govern the transfer of the displacements in the various lattice modes. For $f_{\mathbf{w}jl}$ is the thermal average of $\exp(i\lambda\eta_{\mathbf{w}j})$, where λ is a constant. With rising temperature, the momentum distribution for the oscillator $H_{\mathbf{w}j}$ will tend increasingly to include high as well as low momentum components, and the thermal average of $\exp(i\lambda\eta_{\mathbf{w}j})$ may therefore be expected to decrease with temperature. In fact, at $T=\infty$, when the momentum distribution will be completely random, this thermal average will clearly be zero.

To sum up, we may say that the random thermal motion of the ions opposes the transfer of their mean positions (i.e. $f_{\mathbf{w}jl}$ decreases with T), and, as a result of this, the polaron effective mass increases with temperature.

§ 7. Magnitude of γ_l

By eqns. (6.7) and (6.8),

$$\gamma_l > \gamma_l^0, \qquad (7.1)$$

where

$$\gamma_l^0 = \frac{\nu}{N\hbar} \sum_{\mathbf{w}}' |B_{\mathbf{w}}|^2 \sin^2 \tfrac{1}{2} a w_l \qquad (7.2)$$

is the value of $\gamma_l(\text{opt})$ at $T=0$. The criterion (6.11) will therefore be satisfied if

$$\gamma_l^0 \gg 1.$$

In order to ascertain whether this latter inequality is valid, we require only a calculation which yields a correct order of magnitude for γ_l^0. For this purpose it suffices, using eqn. (7.2), to obtain formula for $B_{\mathbf{w}}$ in terms of $\mathbf{w}$ for the long waves and then to assume that formula to be approximately correct even when $w \sim a^{-1}$.

By (3.11) and (3.15),

$$B_{\mathbf{w}} = \left(\frac{2}{M_0}\right)^{1/2} \nu^{-2} \int |\phi(\mathbf{x})|^2 \sum_t \hat{\mathbf{w}} \exp i\mathbf{w} \cdot \mathbf{a}_t) \nabla [g V(\mathbf{x}-\mathbf{a}_t) - (1-g) V'(\mathbf{x}-\mathbf{a}_t+a\boldsymbol{\alpha})] \, d^3\mathbf{x}.$$

Since, for small $w(\ll a^{-1})$,

$$\sum_{l} \ldots\ldots \exp(i\mathbf{w}\,.\,\mathbf{a}_l) = a^{-3}\int \ldots \exp(i\mathbf{w}\,.\,\mathbf{a}_l)\,d^3\mathbf{a}_l,$$

it follows that we may rewrite the above equation for $B_{\mathbf{w}}$ as

$$B_{\mathbf{w}} = \left(\frac{2}{M_0}\right)^{1/2} \nu^{-2}a^{-3}\int |\phi(\mathbf{x})|^2 \hat{\mathbf{w}} \exp(i\mathbf{w}\,.\,\mathbf{X})\,.\,\frac{\partial}{\partial \mathbf{x}}\,[gV(\mathbf{x}-\mathbf{X}) - (1-g)V'(\mathbf{x}-\mathbf{X}+a\boldsymbol{\alpha})]\,d^3\mathbf{x}\,d^3\mathbf{X}. \quad . \quad . \quad . \quad (7.3)$$

Since, in the atomic state ϕ, the electron is localized within a distance $\ll a \ll w^{-1}$, we may put

$$|\phi(\mathbf{x})|^2 = \delta(\mathbf{x}) \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.4)$$

in our calculation of $B_{\mathbf{w}}$. Also, for small w, we may calculate $B_{\mathbf{w}}$ on the basis of formulae for V and V', which are valid when $|\mathbf{x}-\mathbf{X}|$ is large. Such formulae are

$$\left.\begin{aligned} V(\mathbf{x}-\mathbf{X}) &= -\frac{Ze_0^2}{|\mathbf{x}-\mathbf{X}|} \\ \text{and} \qquad V(\mathbf{x}-\mathbf{X}+a\boldsymbol{\alpha}) &= \frac{Ze_0^2}{|\mathbf{x}-\mathbf{X}+a\boldsymbol{\alpha}|}, \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (7.5)$$

where $\pm Ze_0$ are the changes per positive and negative ion, respectively, $-e_0$ being the electronic charge. Inserting the expressions for $|\phi|^2$, V and V', given by (7.4) and (7.5), into (7.3), we obtain

$$B_{\mathbf{w}} = \left(\frac{2}{M_0}\right)^{1/2} \frac{4\pi e_0^2 Zi}{\nu^2 a^3 w}\,[g + (1-g)\exp(ia\mathbf{w}\,.\,\boldsymbol{\alpha})].$$

For $wa \ll 1$, this reduces to

$$B_{\mathbf{w}} = \left(\frac{2}{M_0}\right)^{1/2} \frac{4\pi e_0^2 Zi}{\nu^2 a^3 w}\,. \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.6)$$

In accordance with the policy stated above, we shall assume this formula to be approximately correct for all the optical modes. Consequently, by (7.2) and (7.6),

$$\gamma_l^0 \simeq \frac{8\pi^2 Z^2 e_0^4}{N\hbar M_0 \nu^3 a^4}\,S_l, \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.7)$$

where

$$S_l = \sum_{\mathbf{w}}{}' \frac{\sin^2 \frac{1}{2}aw_l}{(\frac{1}{2}aw)^2}\,. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.8)$$

Since, by (2.19), $|\frac{1}{2}aw_l| \leqslant \frac{1}{2}\pi$ in the occupied region, it follows that

$$(\tfrac{1}{2}aw_l)^2 \geqslant \sin^2 \tfrac{1}{2}aw_l \geqslant (aw_l/\pi)^2$$

in that region. Consequently, it follows that a lower limit to S_l, giving a correct order of magnitude to that quantity, will be obtained if we replace $\sin^2 \frac{1}{2}aw_l$ by $(aw_l/\pi)^2$ in (7.8). Thus,

$$S_l \simeq \sum_{\mathbf{w}}{}' \frac{4}{\pi^2}\frac{w_l^2}{w^2} = S_l', \text{ say.} \quad . \quad . \quad . \quad . \quad . \quad . \quad (7.9)$$

By symmetry, S_l' is independent of l and is therefore equal to

$$\tfrac{1}{3}\sum_{l=1}^{3} S_l' = \sum_{\mathbf{w}}{}' \frac{4}{3\pi^2}.$$

Hence, by (7.9),

$$S_l \simeq \sum_{\mathbf{w}}{}' \frac{4}{3\pi^2}.$$

The summation in this equation is taken over half of the occupied region in $\mathbf{w}$-space, and therefore

$$S_l \simeq \frac{2}{3\pi^2} N.$$

Inserting this value for S_l into (7.7), we obtain

$$\gamma_l{}^0 \simeq \frac{16 Z^2 e_0{}^4}{3\hbar M_0 \nu^3 a^4}. \qquad \text{(7.10)}$$

Typical values for M_0, a, ν, Z are

$$\left.\begin{aligned} M_0 &= 10^5\, m_{\mathrm{el}}, \\ a &= 5 \times 10^{-8}\ \mathrm{cm}, \\ \nu &= 2 \times 10^{13}\ \mathrm{sec}^{-1}, \\ Z &= 1. \end{aligned}\right\} \qquad \text{(7.11)}$$

The lattice distance a, given here, is about ten Bohr radii, which is sufficiently large for the tight-binding approach to be justified, provided that the ionic radii $\ll a$.

On using the values given by (7.11), eqn. (7.10) yields

$$\gamma_l{}^0 \simeq 60. \qquad \text{(7.12)}$$

Hence, in view of (7.1), (6.11) can be satisfied for reasonable values of M_0, a, ν, Z. This confirms our assertion that the polaron effective mass, given by (6.9), can be extremely large in cases where the tight-binding approach is valid.

In the above calculations, the ultra-violet polarization, due to the deformation of the ions, has been ignored. We can take this polarization into account by dividing the potentials V and V', given by (7.5), by a dielectric constant K. K is in fact the high-frequency dielectric constant, K_∞, since the ultra-violet polarization can follow the field due to the electron adiabatically. It is easily seen that the division of V and V' by K_∞ leads to the division of our expression for $\gamma_l{}^0$ by $K_\infty{}^2$. Consequently, instead of (7.12) we have

$$\gamma_l{}^0 \simeq 60/K_\infty{}^2.$$

With $K_\infty = 2$ (a typical value for a polar crystal), we then have $\gamma_l{}^0 \simeq 15$. This value for $\gamma_l{}^0$ is still large enough for its exponential to be extraordinarily large. Therefore our main thesis, that the polaron effective mass may be extraordinarily large in cases where the tight-binding

approach is valid, is seen to still hold good, even when the screening effects due to the ultra-violet polarization are taken into account.

APPENDIX 1

We may rewrite eqn. (6.4) as

$$f_{\mathbf{w}jt}=\frac{Sp[\exp(i\lambda_{\mathbf{w}j}\eta_{\mathbf{w}j})\exp(-\beta H_{\mathbf{w}j})]}{Sp[\exp(-\beta H_{\mathbf{w}j})]}, \quad . \quad . \quad . \quad \text{(A1.1)}$$

where

$$\lambda_{\mathbf{w}j}=(\bar{\xi}_{\mathbf{w}j0}-\bar{\xi}_{\mathbf{w}jt})/\hbar. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(A1.2)}$$

It may be seen from (2.22) that $H_{\mathbf{w}j}$ is unchanged by the canonical transformation

$$\eta_{\mathbf{w}j}\rightarrow\nu_{\mathbf{w}j}\xi_{\mathbf{w}j}, \quad \xi_{\mathbf{w}j}\rightarrow-\eta_{\mathbf{w}j}/\nu_{\mathbf{w}j}.$$

Consequently, we may replace $H_{\mathbf{w}j}$, $\eta_{\mathbf{w}j}$, $\xi_{\mathbf{w}j}$ by $H_{\mathbf{w}j}$, $\nu_{\mathbf{w}j}\xi_{\mathbf{w}j}$, $-\eta_{\mathbf{w}j}/\nu_{\mathbf{w}j}$, respectively, in (A1.1). Hence

$$f_{\mathbf{w}jt}=\frac{Sp[\exp(i\lambda_{\mathbf{w}j}\nu_{\mathbf{w}j}\xi_{\mathbf{w}j})\exp(-\beta H_{\mathbf{w}j})]}{Sp[\exp(-\beta H_{\mathbf{w}j})]}. \quad . \quad . \quad \text{(A1.3)}$$

The eigenstates and eigenvalues of $H_{\mathbf{w}j}$ are $\theta(n_{\mathbf{w}j}, \xi_{\mathbf{w}j})$ and $(\eta_{\mathbf{w}j}+\frac{1}{2})\hbar\nu_{\mathbf{w}j}$, respectively, where the θ's are real, normalized functions. Therefore, if we drop the suffixes on λ, ν, ξ and n, we may rewrite (A1.3) as

$$f_{\mathbf{w}jt}=\int_{-\infty}^{\infty}\rho(\xi)\exp(i\lambda\nu\xi)\,d\xi, \quad . \quad . \quad . \quad . \quad \text{(A1.4)}$$

where

$$\rho(\xi)=\frac{\sum_{n=0}^{\infty}[\theta(n,\xi)]^2\exp(-n\beta\hbar\nu)}{\sum_{n=0}^{\infty}\exp(-n\beta\hbar\nu)}. \quad . \quad . \quad . \quad . \quad . \quad \text{(A1.5)}$$

In order to evaluate $\rho(\xi)$, we make use of the well known relations (cf. Bohm 1954, p. 304), for the eigenstates of a harmonic oscillator, namely

$$\left.\begin{aligned}&\left(\nu\xi-\hbar\frac{\partial}{\partial\xi}\right)\theta(n,\xi)=[2\hbar(n+1)]^{1/2}\theta(n+1,\xi)\\ \text{and}\\ &\left(\nu\xi+\hbar\frac{\partial}{\partial\xi}\right)\theta(n,\xi)=[2\hbar n]^{1/2}\theta(n-1,\xi).\end{aligned}\right\} \quad \text{(A1.6)}$$

It now follows from (A1.5), (A1.6) that

$$\left(\nu\xi-\frac{1}{2}\hbar\frac{\partial}{\partial\xi}\right)\rho=C\sum_{n=0}^{\infty}(n+1)^{1/2}\exp(-n\beta\hbar\nu)\theta(n,\xi)\theta(n+1,\xi) \quad \text{(A1.7)}$$

and

$$\left(\nu\xi+\frac{1}{2}\hbar\frac{\partial}{\partial\xi}\right)\rho=C\sum_{n=0}^{\infty}n^{1/2}\exp(-n\beta\hbar\nu)\theta(n-1,\xi)\theta(n,\xi), \quad \text{(A1.8)}$$

where

$$C=(2\hbar)^{1/2}\left[\sum_{n=0}^{\infty}\exp(-n\beta\hbar\nu)\right]^{-1}.$$

Equation (A1.8) may be rewritten as

$$\left(\nu\xi + \tfrac{1}{2}\hbar\frac{\partial}{\partial\xi}\right)\rho = C\sum_{n=1}^{\infty} n^{1/2}\exp\left(-n\beta\hbar\nu\right)\theta(n-1,\,\xi)\theta(n,\,\xi),$$

since the summand is zero for $n=0$. On replacing $(n-1)$ by n in this last equation, we obtain

$$\left(\nu\xi + \tfrac{1}{2}\hbar\frac{\partial}{\partial\xi}\right)\rho = C\sum_{n=0}^{\infty}(n+1)^{1/2}\exp\left[-(n+1)\beta\hbar\nu\right]\theta(n,\,\xi)\theta(n+1,\,\xi).$$

Consequently, by (A1.7),

$$\left(\nu\xi + \tfrac{1}{2}\hbar\frac{\partial}{\partial\xi}\right)\rho = \exp\left(-\beta\hbar\nu\right)\left(\nu\xi - \tfrac{1}{2}\hbar\frac{\partial}{\partial\xi}\right)\rho,$$

i.e.

$$\frac{\partial\rho}{\partial\xi} = -\frac{2\nu\xi}{\hbar}\tanh\left(\tfrac{1}{2}\beta\hbar\nu\right)\rho,$$

from which it follows that

$$\rho = \rho_0\exp\left[-\frac{\nu\xi^2}{\hbar}\tanh\left(\tfrac{1}{2}\beta\hbar\nu\right)\right], \quad . \quad . \quad . \quad . \quad \text{(A1.9)}$$

where ρ_0 is independent of ξ and λ. It now follows from (A1.4) and (A1.9) that

$$f_{\mathbf{w}jt} = A\exp\left[-\tfrac{1}{4}\hbar\nu\lambda^2\coth\left(\tfrac{1}{2}\beta\hbar\nu\right)\right], \quad . \quad . \quad . \quad \text{(A1.10)}$$

where

$$A = \rho_0\int_{-\infty}^{\infty}\exp\left[-\frac{\nu}{\hbar}\tanh\left(\tfrac{1}{2}\beta\hbar\nu\right)\xi^2\right]d\xi.$$

A is therefore independent of λ. By eqn. (A1.3), $f_{\mathbf{w}jt}=1$ when $\lambda=0$. It follows from (A1.10) that $A=1$. Hence, on reasserting the suffixes to ν and λ, we may rewrite (A1.10) as

$$f_{\mathbf{w}jt} = \exp\left[-\tfrac{1}{4}\hbar\nu_{\mathbf{w}j}\lambda_{\mathbf{w}j}{}^2\coth\left(\tfrac{1}{2}\beta\hbar\nu_{\mathbf{w}j}\right)\right]. \quad . \quad . \quad . \quad \text{(A1.11)}$$

Equation (6.5) follows immediately on substituting into (A1.11) the value for $\lambda_{\mathbf{w}j}$ given by (A1.1).

APPENDIX 2

In order to prove eqns. (5.6), we note that, by (5.3),

$$W^{+}(\mathbf{a}_l) = W(-\mathbf{a}_l), \quad . \quad . \quad . \quad . \quad . \quad \text{(A2.1)}$$

since $\epsilon(\mathbf{k})$ is real.

By eqn. (4.13), $W^0(\mathbf{a}_l)$ is real, since the atomic wave-function ϕ is real, and, by (6.3) and (6.5), $F(\mathbf{a}_l)$ is real. Hence, by (5.4), $W(\mathbf{a}_l)$ is real, so that

$$W^{+}(\mathbf{a}_l) = W(\mathbf{a}_l).$$

Equation (5.6) now follows from this last eqn. and (2.1).

APPENDIX 3

We wish to show that, for given $\mathbf{k}$, the mean square deviation from $\epsilon(\mathbf{k})$ of the energies $\epsilon_{\mathbf{k}n}$, at temperature T, is negligible. This deviation is the difference between the thermal average, taken over lattice states, of

$\epsilon_{\mathbf{k}n}{}^2$, and the square of the average of $\epsilon_{\mathbf{k}n}$. It may be seen from (4.14) and (4.15) that this deviation is negligible if the mean square deviation of F_n is negligible, where

$$F_n = (\Theta_n{}^+ | T_0{}^+ T_t | \Theta_n). \quad \text{(A3.1)}$$

This latter mean square deviation is

$$\Delta = \overline{F}^2 - \overline{F}^2, \quad \text{(A3.2)}$$

where

$$\overline{F}^r = \frac{\sum_n F_n{}^r \exp(-\beta E_n(\text{lat}))}{\sum_n \exp(-\beta E_n(\text{lat}))}, \quad \text{for } r = 1, 2, \ldots \quad \text{(A3.3)}$$

so that, by (5.5), (A3.1) and (A3.3),

$$\overline{F} = F(\mathbf{a}_t). \quad \text{(A3.4)}$$

It remains for us to prove that Δ is negligibly small.

By eqn. (A3.2),

$$\frac{\Delta}{\overline{F}^2} = \frac{\overline{F}^2}{\overline{F}^2} - 1;$$

and hence, since $\overline{F}^2 < 1$, as may be seen from (6.3), (6.5) and (A3.4), it follows that

$$\Delta < \frac{\overline{F^2}}{\overline{F}^2} - 1; \quad \text{(A3.5)}$$

In order to evaluate the r.h.s. of this inequality, we note that

$$(\theta^+(n, \xi) | \exp(i\lambda\eta) | \theta(n, \xi)) = (\theta^+(n, \xi) | \cos\lambda\eta | \theta(n, \xi)), \quad \text{(A3.6)}$$

since the expectation value, in the state $\theta(n, \xi)$, of any odd power of the momentum η is zero. It now follows from (2.25), (3.17), (A1.2), (A3.1) and (A3.6) that

$$F_n = (\Theta_n{}^+ | G | \Theta_n), \quad \text{(A3.7)}$$

where

$$G = \prod_{j=1}^{4} \prod_{\mathbf{w}}{}' \cos\lambda_{\mathbf{w}j}\eta_{\mathbf{w}j}. \quad \text{(A3.8)}$$

The quantity

$$(\Theta_n{}^+ | G^2 | \Theta_n) - (\Theta_n{}^+ | G | \Theta_n)^2$$

is the mean square deviation of G, for the state Θ_n, and is therefore positive. Consequently, by (A3.7),

$$F_n{}^2 < (\Theta_n{}^+ | G^2 | \Theta_n). \quad \text{(A3.9)}$$

It now follows from eqns. (2.21), (A3.3), (A3.7) and (A3.9) that

$$\overline{F} = \prod_{j=1}^{4} \prod_{\mathbf{w}}{}' \text{av}(\cos\lambda_{\mathbf{w}j}\eta_{\mathbf{w}j}) \quad \text{(A3.10)}$$

and

$$\overline{F}^2 < \prod_{j=1}^{4} \prod_{\mathbf{w}}{}' \text{av}(\cos^2\lambda_{\mathbf{w}j}\eta_{\mathbf{w}j}), \quad \text{(A3.11)}$$

where, for any function $g(\eta)$,

$$\mathrm{av}(g(\eta_{\mathbf{w}j}))\equiv \frac{Sp[g(\eta_{\mathbf{w}j})\exp(-\beta H_{\mathbf{w}j})]}{Sp[\exp(-\beta H_{\mathbf{w}j})]} \quad . \quad . \quad . \quad . \quad \text{(A3.12)}$$

Consequently, by (A3.5), (A3.10) and (A3.11),

$$1+\Delta < \prod_{j=1}^{4}\prod_{\mathbf{w}}{}' \frac{\mathrm{av}(1+\cos 2\lambda_{\mathbf{w}j}\eta_{\mathbf{w}j})}{2\mathrm{av}(\cos\lambda_{\mathbf{w}j}\eta_{\mathbf{w}j})} \quad . \quad . \quad . \quad . \quad \text{(A3.13)}$$

Moreover, by eqns. (A1.1) and (A3.6), $f_{\mathbf{w}jl}$ is the thermal average of $\cos(\lambda_{\mathbf{w}j}\eta_{\mathbf{w}j})$, for the oscillator $H_{\mathbf{w}j}$, and therefore, by (A3.12),

$$f_{\mathbf{w}jl}=\mathrm{av}(\cos\lambda_{\mathbf{w}j}\eta_{\mathbf{w}j}).$$

Hence, by (A1.11),

$$\mathrm{av}(\cos\lambda_{\mathbf{w}j}\eta_{\mathbf{w}j})=\exp(-\tfrac{1}{2}\sigma_{\mathbf{w}j}), \quad . \quad . \quad . \quad \text{(A3.14)}$$

where

$$\sigma_{\mathbf{w}j}=\tfrac{1}{2}\hbar\nu_{\mathbf{w}j}\lambda^2{}_{\mathbf{w}j}\coth(\tfrac{1}{2}\beta\hbar\nu_{\mathbf{w}j}). \quad . \quad . \quad . \quad . \quad \text{(A3.15)}$$

It follows from (A3.15) that we may replace λ, σ by 2λ, 4σ, respectively, in (A3.14). Hence

$$\mathrm{av}(\cos 2\lambda_{\mathbf{w}j}\eta_{\mathbf{w}j})=\exp(-2\sigma_{\mathbf{w}j});$$

and therefore, by (A3.13), (A3.14),

$$1+\Delta < \prod_{j=1}^{4}\prod_{\mathbf{w}}{}' \cosh\sigma_{\mathbf{w}j}. \quad . \quad . \quad . \quad . \quad . \quad \text{(A3.16)}$$

In order to simplify this inequality, we note that, since $t>\tanh t$, for $t>0$,

$$0<\int_0^{\sigma}(t-\tanh t)\,dt=\tfrac{1}{2}\sigma^2-\ln\cosh\sigma;$$

and therefore, $\cosh\sigma<\exp(\frac{1}{2}\sigma^2)$. Inserting this inequality into (A3.16),

$$\Delta<\exp(\Lambda)-1, \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(A3.17)}$$

where

$$\Lambda=\sum_{j=1}^{4}\sum_{\mathbf{w}}{}'\tfrac{1}{2}\sigma_{\mathbf{w}j}{}^2.$$

It follows from this last equation that, since the σ's are positive,

$$\Lambda<\tfrac{1}{2}\sum_{\mathbf{w}}{}'\tau_{\mathbf{w}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(A3.18)}$$

where

$$\tau_{\mathbf{w}}=(\sigma_{\mathbf{w}1}+\sigma_{\mathbf{w}2})^2+(\sigma_{\mathbf{w}3}+\sigma_{\mathbf{w}4})^2. \quad . \quad . \quad . \quad . \quad \text{(A3.19)}$$

Since τ is an even function of $\mathbf{w}$, as may be seen from (3.14), (3.15), (A1.2), (A3.15) and (A1.19), we may replace $\sum_{\mathbf{w}}{}'$ by $\frac{1}{2}\sum_{\mathbf{w}}$ in (A3.18). Hence,

$$\Lambda<\tfrac{1}{4}\sum_{\mathbf{w}}\tau_{\mathbf{w}}. \quad . \quad . \quad . \quad . \quad . \quad \text{(A3.20)}$$

On expressing $\tau_{\mathbf{w}}$ in terms of $B_{\mathbf{w}}$, $C_{\mathbf{w}}$, ν and c, by means of (2.24), (3.14), (A1.2), (A3.15) and (A3.19), and on replacing

$$\sum_{\mathbf{w}}\ldots\ldots \quad \text{by} \quad \frac{Na^3}{(2\pi)^3}\int\ldots d^3\mathbf{w},$$

(A3.20) yields

$$\Lambda < N^{-1}I, \qquad \text{(A3.21)}$$

where

$$I = \frac{a^3}{(2\pi)^3 h^2}\int [|B_{\mathbf{w}}|^4 \nu^2 \coth^2(\tfrac{1}{2}\beta h\nu) + |C_{\mathbf{w}}|^4 c^2 w^2 \coth^2(\tfrac{1}{2}\beta hcw)] \sin^4 \tfrac{1}{2}\mathbf{w}\,.\,\mathbf{a}_l\, d^3\mathbf{w}. \qquad \text{(A3.22)}$$

The integration in this last equation is taken over the whole of the occupied region of $\mathbf{w}$-space. It follows from (3.15) and (A3.22) that I is independent of N. Moreover, the contributions to the integral in (A3.22), due to the long waves, do not lead to a divergence. This may easily be verified on evaluating $B_{\mathbf{w}}$ and $C_{\mathbf{w}}$ for the long waves, using the method employed in § 7 to evaluate $B_{\mathbf{w}}$. I is therefore finite.

It follows from (A3.21) that

$$\Lambda \to 0 \quad \text{as} \quad N \to \infty$$

and therefore, by (A3.17), $\Delta \to 0$ as $N \to \infty$. Consequently, Δ is negligibly small for a macroscopic body.

Acknowledgment

The author would like to express his thanks to Professor H. Fröhlich, F.R.S., for suggesting this investigation and also for a number of very helpful discussions.

References

Allcock, G. R., 1956, *Advanc. Phys.*, **5,** 412.
Bloch, F., 1928, *Z. Phys.*, **52,** 555.
Bohm, D., 1954, *Quantum Theory* (London: Constable and Co.).
Born, M., and Huang, K., 1954, *Dynamical Theory of Crystal Lattices* (Oxford).
Fröhlich, H., 1954, *Advanc. Phys.*, **3,** 325; 1957, *Arch. Sci., Genève,* **10,** 5.
Morin, F. J., 1954, *Phys. Rev.*, **93,** 1195, 1199.
Wilson, A. H., 1953, *Theory of Metals* (Cambridge).

Calculation of the Thermal Expansion of Solids from the Third-Order Elastic Constants†

By F. W. Sheard
Cavendish Laboratory, Cambridge

[Received July 20, 1958]

Abstract

The thermal expansion of a solid is due to anharmonicity of the interatomic forces. It may therefore be calculated, on the basis of a continuum model, from the observed elastic anharmonicity, i.e. the third-order elastic constants. Using Lazarus' (1949) measurements on the pressure variation of the elastic moduli of several cubic crystals, the Grüneisen γ is calculated for the limiting cases of low and high temperatures. For KCl, NaCl the high temperature values agree with those obtained from thermal expansivities, and a decrease in γ at low temperatures is predicted. However, Lazarus' data on metals cannot be reconciled with the thermal expansions, which suggests that the continuum model is inapplicable to these cases.

§ 1. Introduction

The assumptions of the Debye theory of specific heats are that the normal modes of vibration of the solid are independent, and that the frequencies of the modes, up to the limiting frequency, are those of a perfectly elastic, isotropic continuum. In this theory everything depends on one parameter, the Debye temperature Θ, which may be calculated from a knowledge of the atomic volume and macroscopic elastic constants of the solid.

This theory is inadequate to describe such phenomena as thermal expansion and the lattice thermal conductivity which depend on the anharmonic nature of the interatomic forces. Calculations of the lattice conductivity have been carried out (Leibfried and Schlömann 1954, Klemens 1956) and have relied upon a single phenomenological parameter as an average measure of the anharmonicity, namely the Grüneisen constant γ (Grüneisen 1926). This is defined experimentally by $\gamma = \alpha V/\chi_T C_V$, where α is the thermal expansivity, χ_T the isothermal compressibility and C_V the heat capacity at constant volume V.

In the Grüneisen theory γ is a measure of the change of the lattice frequencies with the volume of the crystal. This may be described in the continuum model by the addition of terms of the third-order in the strain components to the elastic strain energy. Since elastic anharmonicity has been studied experimentally it is interesting to see whether from such information we can predict the observed values of γ obtained from thermal expansion measurements.

† Communicated by Dr. J. M. Ziman.

Bridgman's measurements on the change of compressibility with pressure, were used in this way by Slater (1939), and fair agreement between the calculated and experimental values of γ was obtained. However, more detailed information on elastic anharmonicity has been given by Lazarus (1949) who measured the pressure variation of the individual elastic moduli of several cubic crystals: in this paper we employ these data to calculate γ on the basis of the anisotropic continuum model. It must be emphasized that the values obtained have only the limited validity of the continuum model.

We shall first derive the theoretical formulae needed for the calculation.

§ 2. Approximate Theoretical Formulae

Consider only monatomic, cubic crystals: the normal modes of vibration may be labelled by a wave vector $\mathbf{q}$ and polarization index p; $p=1, 2, 3$ corresponding to one longitudinal and two transverse branches of the frequency spectrum. By treating the crystal as an assembly of loosely coupled harmonic oscillators, whose frequencies depend on the volume of the crystal, and applying statistical mechanics, it is easy to calculate the free energy F from which the Grüneisen parameter follows as a consequence of the thermodynamic relation

$$\frac{\alpha}{\chi_T} = -\left(\frac{\partial^2 F}{\partial V \partial T}\right).$$

The result is (Slater 1939, Barron 1955):

$$\gamma = \sum_{\mathbf{q},p} \gamma_{\mathbf{q},p} C_{\mathbf{q},p} \Big/ \sum_{\mathbf{q},p} C_{\mathbf{q},p}, \qquad (1)$$

where

$$C_{\mathbf{q},p} = k\left(\frac{\hbar\omega_{\mathbf{q},p}}{kT}\right)^2 \frac{\exp(\hbar\omega_{\mathbf{q},p}/kT)}{\{\exp(\hbar\omega_{\mathbf{q},p}/kT)-1\}^2}, \qquad (2)$$

and

$$\gamma_{\mathbf{q},p} = -\left(\frac{\partial \ln \omega_{\mathbf{q},p}}{\partial \ln V}\right)_T; \qquad (3)$$

$\omega_{\mathbf{q},p}$ is the frequency of a normal mode and the sums are over the $3N$ possible modes (N is the number of unit cells). Thus γ is expressed as a weighted average of the individual $\gamma_{\mathbf{q},p}$'s. The weighting factor $C_{\mathbf{q},p}$ is, of course, the contribution of the $(\mathbf{q},p)$th mode to the heat capacity $C_V = \sum_{\mathbf{q},p} C_{\mathbf{q},p}$.

The simplifying property of a continuum is the absence of dispersion of the elastic waves. The frequencies are given by $\omega_{\mathbf{q},p} = q s_p(\theta,\phi)$ where $q = |\mathbf{q}|$ and $s_p(\theta,\phi)$ is the sound velocity of the pth branch in the direction specified by polar angles θ, ϕ. Hence

$$\gamma_{\mathbf{q},p} = -\frac{d\ln q}{d\ln V} - \left(\frac{\partial \ln s_p(\theta,\phi)}{\partial \ln V}\right)_T.$$

The wave vector $\mathbf{q}$, according to periodic boundary conditions applied to a cube of side L, has components $2\pi n_i/L$ (integer values of n_i); thus for a particular mode, the n_i are fixed and $q \propto V^{-1/3}$. We can write $\gamma_{\mathbf{q},p} = \gamma_p(\theta,\phi)$, where

$$\gamma_p(\theta,\phi) = \frac{1}{3} - \left(\frac{\partial \ln s_p(\theta,\phi)}{\partial \ln V}\right)_T . \qquad (4)$$

Equation (1) can be considerably simplified for the limiting cases of low and high temperatures.

For low temperatures, $T \ll \Theta$, converting the sums to integrals over $\mathbf{q}$-space and using standard Debye theory, we obtain

$$\gamma = \sum_p \int \frac{\gamma_p(\theta,\phi)}{\{s_p(\theta,\phi)\}^3}\, d\Omega \Big/ \sum_p \int \frac{d\Omega}{\{s_p(\theta,\phi)\}^3}, \qquad (5)$$

where $d\Omega$ is an element of solid angle in the direction θ,ϕ.

For high temperatures, $T > \Theta$, the heat capacity per mode $C_{\mathbf{q},p} = k$ and hence

$$\gamma = \frac{1}{3N} \sum_{\mathbf{q},p} \gamma_{\mathbf{q},p},$$

which, in the continuum approximation becomes

$$\gamma = \frac{1}{12\pi} \sum_p \int \gamma_p(\theta,\phi)\, d\Omega. \qquad (6)$$

The experimental information we shall use refers to room temperatures, for which $T \sim \Theta$. Writing $C(x) = C_{\mathbf{q},p}/k = x^2e^x/(e^x-1)^2$, and remembering the limiting frequency for the longitudinal branch is about twice that for a transverse branch, we have at room temperature

$$0 < x < \frac{\hbar\omega_{\max}}{kT} \sim \frac{\hbar\omega_{\max}}{k\Theta} \sim \begin{cases} 0{\cdot}5 \text{ for transverse branch,} \\ 1{\cdot}5 \text{ for longitudinal branch.} \end{cases}$$

Hence $1 > C(x) > \begin{cases} 0{\cdot}98 \\ 0{\cdot}83 \end{cases}$; that is, the weighting factor for the higher frequencies of either branch has not fallen significantly below the upper limit, unity. Equation (6) is therefore a good approximation at room temperatures.

The difference between the low and high temperature eqns. (5) and (6) respectively, lies simply in the type of average taken over the $\gamma_p(\theta,\phi)$. These limiting values of γ will be denoted by γ_{L}, γ_{H}.

§ 3. Anharmonicity in an Elastic Continuum

Before explicitly evaluating $\gamma_p(\theta,\phi)$ we shall consider briefly how the anharmonicity is specified in terms of the third-order constants, and the available experimental information concerning them.

The free energy per unit (unstrained) volume may be written as an expansion in the Lagrangian strain components η_r, at a fixed temperature:

$$f=\frac{1}{2}\sum_{p\leqslant q} c_{pq}\eta_p\eta_q+\sum_{p\leqslant q\leqslant r} C_{qpr}\eta_p\eta_q\eta_r, \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where p, q, r take values 1, 2, 3 . . . 6 (Birch 1947). The second-order coefficients are the elastic moduli of Voigt, and the C_{pqr} are the third-order or anharmonic coefficients. For cubic crystals belonging to the classes of maximum symmetry there are only three independent elastic constants c_{11}, c_{12}, c_{44} and six independent third-order constants $C_{111}, C_{112}, C_{144}, C_{166}, C_{123}, C_{456}$.

Since cubic symmetry is preserved under uniform pressure the form of the relations between an additional infinitesimal stress and corresponding infinitesimal strain are also preserved, but the 'effective elastic constants' depend on the pressure. Birch (1947) showed that the elastic moduli under a finite pressure, corresponding to volume V, are

$$\left.\begin{aligned} c_{11}(V)&=c_{11}+\frac{\Delta V}{3V_0}(2c_{11}+2c_{12}+6C_{111}+4C_{112}),\\ c_{12}(V)&=c_{12}+\frac{\Delta V}{3V_0}(C_{123}+4C_{112}-c_{11}-c_{12}),\\ c_{44}(V)&=c_{44}+\frac{\Delta V}{3V_0}(c_{44}+c_{11}+2c_{12}+\tfrac{1}{2}C_{144}+C_{166}),\end{aligned}\right\} \quad . \quad . \quad . \quad (8)$$

to the first order in $\Delta V=V-V_0$. At zero pressure, $V=V_0$, these reduce to c_{11}, c_{12}, c_{44}, i.e. the coefficients in eqn. (7).

Lazarus (1949) has determined the adiabatic elastic constants of several cubic crystals by measuring the velocities of ultrasonic waves up to pressures of 10 000 bars, and Hearmon (1953), by plotting the dependence upon volume, deduced values for the quantities

$$C_a=6C_{111}+4C_{112}, \quad C_b=C_{123}+4C_{112}, \quad C_d=\tfrac{1}{2}C_{144}+C_{166}, \quad . \quad . \quad (9)$$

using the above relations. It is clear that in the derivation of the latter equations the $c_{11}(V)$ etc. are taken to be the isothermal moduli and the volume change ΔV is carried out isothermally. Since Hearmon omitted to convert Lazarus' adiabatic moduli to the corresponding isothermal moduli before applying eqns. (8), his values for the third-order constants are incorrect. The errors are not likely to be great; in the case of Cu the corrections to the values of C_a, C_b, C_d are estimated to be less than 10%.

However, for our calculation of the Grüneisen ratio we require the dependence of sound velocities upon isothermal volume changes and this will be given correctly by substituting Lazarus' values for the adiabatic moduli (distinguished by a superscript s) and Hearmon's values for C_a, C_b, C_d (also distinguished by a superscript s) into the equations:

$$c_{11}^{s}(V)=c_{11}^{s}+\frac{\Delta V}{3V_0}(2c_{11}^{s}+2c_{12}^{s}+C_a^{s}), \text{ etc.} \quad . \quad . \quad . \quad (10)$$

§ 4. The Magnitude of γ at Room Temperature

Previous estimations of γ (Slater 1939) neglected crystal anisotropy and relied on special assumptions, for example that Poisson's ratio is independent of volume. We can avoid these approximations by obtaining a general expression for $\gamma_p(\theta, \phi)$ as follows.

The secular equation giving the sound velocity $s(\theta, \phi)$ in a cubic continuum may be written (Quimby and Sutton 1953):

$$z^3 - z^2 + B_1\Gamma_1(\theta,\phi)z - B_2\Gamma_2(\theta,\phi) = 0; \quad \ldots \quad (11)$$

the variable z is related to s by

$$z = \rho s^2/K_3 - K_1 \quad \ldots \quad (12)$$

and

$$\left.\begin{aligned} &B_1 = 1 - K_2^2, \quad B_2 = 1 - 3K_2^2 + 2K_2^3, \\ &K_1 = c_{44}^s/K_3, \quad K_2 = (c_{12}^s + c_{44}^s)/K_3, \quad K_3 = c_{11}^s - c_{44}^s, \end{aligned}\right\} \quad (13)$$

$$\Gamma_1(\theta,\phi) = \sin^2\theta(\sin^2\theta\sin^2\phi\cos^2\phi + \cos^2\theta),$$

$$\Gamma_2(\theta,\phi) = \sin^4\theta\cos^2\theta\sin^2\phi\cos^2\phi.$$

The elastic constants appearing here refer to the particular pressure at which we require the sound velocities; they are the $c_{11}^s(V)$ of eqn. (10).

Differentiating eqns. (11) and (12) with respect to $\ln V$ at constant temperature and eliminating $(\partial z/\partial \ln V)_T$ immediately gives $(\partial \ln s/\partial \ln V)_T$;

$$2\left(\frac{\partial \ln s(\theta,\phi)}{\partial \ln V}\right)_T = \frac{(B_2'\Gamma_2 - zB_1'\Gamma_1)}{(3z^2 - z + B_1\Gamma_1)(z + K_1)} + \frac{K_1'}{z + K_1} + \frac{K_3'}{K_3} + 1, \quad (14)$$

where we have used the notation $F'(V) \equiv (\partial F(V)/\partial \ln V)_T$, for any function of V.

The quantities K_i', B_i' can be expressed in terms of the elastic moduli and their variation with volume. For example:

$$\left.\begin{aligned} &c_{11}^{s\prime} = \tfrac{1}{3}(2c_{11}^s + 2c_{12}^s + C_a^s), \quad \text{at } V = V_0, \\ &K_2' = (c_{11}^s c_{12}^{s\prime} - c_{11}^{s\prime} c_{12}^s + c_{12}^s c_{44}^{s\prime} - c_{12}^{s\prime} c_{44}^s + c_{11}^s c_{44}^{s\prime} - c_{11}^{s\prime} c_{44}^s)/K_3^2, \\ &B_1' = -6K_2K_2'. \end{aligned}\right\} \quad (15)$$

For special directions eqn. (11) can easily be solved for the sound velocities (de Launay 1956) whence eqn. (14) gives the values of γ_p, e.g. in the [100] direction

$$\left.\begin{aligned} &s_1 = (c_{11}^s/\rho)^{1/2}; \quad s_{2,3} = (c_{44}^s/\rho)^{1/2}; \\ &\left(\frac{\partial \ln s}{\partial \ln V}\right)_T = \tfrac{1}{2}(c_{11}^{s\prime}/c_{11}^s + 1); \quad \tfrac{1}{2}(c_{44}^{s\prime}/c_{44}^s + 1), \quad \text{twice.} \end{aligned}\right\} \quad (16)$$

To calculate γ for a given substance we take from experiment the values of ρ, c_{11}^s, c_{12}^s, c_{44}^s, C_a^s, C_b^s, C_d^s. Equations (12), (13) and (15) at once give us $c_{11}^{s\prime} \ldots K_1 \ldots B_2$, $K_1' \ldots B_2'$. For a particular direction θ, ϕ the secular equation is solved, the three roots z_p give the sound velocities s_p

and substituting in eqns. (14) and (4) the values of γ_p for this direction are obtained. Numerical integration according to low or high temperature equations yields γ.

The integrals were evaluated numerically on EDSAC 2, the computer at the University Mathematical Laboratory, a Gauss twelve-point integration formula being used over the range $\theta, \phi = 0°$ to $90°$. It was found that eqn. (14) is indeterminate for the longitudinal branch ($p = 1$) in the |111| direction (it gives 0/0), but this does not invalidate the integration procedure as the points at which the integrand is evaluated do not approach closer than 6° to this special direction.

Since in the course of the calculation a mean sound velocity s_m and Debye temperature Θ are evaluated, they are recorded below for their own intrinsic interest. Their definitions are

$$\frac{1}{{s_m}^3} = \frac{1}{12\pi} \sum_{p=1}^{3} \int \frac{d\Omega}{\{s_p(\theta, \phi)\}^3}, \qquad \Theta = \frac{h}{k} \left(\frac{6\pi^2 N n}{V} \right)^{1/3} s_m. \quad . \quad . \quad (17)$$

(n is the number of atoms per unit cell, N the number of unit cells.)

To see the contributions from the individual polarization branches to the final average γ and also the effect of anisotropy of the medium, the sound velocities s_p and Grüneisen constants γ_p are recorded for the [100], [110] and [111] directions using (16) and similar equations.

The experimental values and data from which they were evaluated are included in table 1.

Table 1

Material	α_l (10^{-5} °K^{-1})	C_p (cal. mole^{-1} °K^{-1})	V (cm^3)	χ_s (cm^2 dyne^{-1} 10^{-12})	$\gamma = \frac{3\alpha_l V}{\chi_s C_p}$
Cu	1·67[a]	5·848[e]	7·11	0·716[i]	2·03
Al	2·36[a]	5·817[e]	9·99	1·285[i]	2·26
KCl	3·67[b]	12·21[f]	37·54	5·450[i]	1·48
NaCl	3·98[c]	11·86[g]	27·04	4·076[i]	1·60
CuZn	2·00[d]	11·61[h]	15·54	0·861[i]	2·22

α_l = coefficient of linear expansion
C_p = heat capacity per mole at constant pressure
V = volume per mole
χ_s = adiabatic compressibility

} at 298°K and atmospheric pressure.

a: Bijl and Pullan (1955).
b: Dayal (1944).
c: Buffington and Lattimer (1926).
d: International Critical Tables.
e: Giauque and Meads (1941).
f: Berg and Morrison (1957).
g: Clusius *et al.* (1949).
h: Sykes and Wilkinson (1937).
i: Lazarus (1949).

Table 2

Material	Direction	$s_1 : L$ †	$s_2 : T$	$s_3 : T$	$\gamma_1 : L$	$\gamma_2 : T$	$\gamma_3 : T$	γ_L	γ_H	s_m †	Θ (°K)
Cu	100	4·37	2·88	2·88	1·69	0·59	0·59	0·762	0·937	2·537	331·7
	110	5·00	2·88	1·62	1·41	0·59	0·80				
	111	5·18	2·12	2·12	1·40	0·69	0·69				
Al	100	6·25	3·25	3·25	2·84	2·57	2·57	3·77	3·26	3·430	400·2
	110	6·41	3·25	3·93	2·28	2·57	5·48				
	111	6·57	2·96	2·96	2·12	4·23	4·23				
KCl	100	4·54	1·78	1·78	2·73	−0·87	−0·87	0·525	1·37	2·377	225·4
	110	3·90	1·78	2·92	1·81	−0·87	3·03				
	111	3·68	2·60	2·60	1·35	2·41	2·41				
NaCl	100	4·80	2·44	2·44	2·81	0·14	0·14	1·23	1·60	2·900	307·0
	110	4·52	2·44	2·92	1·99	0·14	2·89				
	111	4·40	2·76	2·76	1·70	2·18	2·18				
CuZn	100	3·94	3·15	3·15	1·90	1·69	1·69	1·74	1·76	2·201	280·3
	110	4·93	3·15	3·42	1·82	1·69	1·75				
	111	5·22	2·02	2·02	1·80	1·70	1·70				

L : longitudinal branch; T : transverse branch.
† In 10^5 cm sec^{-1}.

The Θ values agree well with other approximate calculations (Blackman 1955) but the values of γ are significantly different. For comparison we have calculated γ using Slater's theory which, in addition to the approximations already mentioned, neglects the difference between the adiabatic and isothermal compressibilities. The resulting formula is

$$\gamma = -2/3 + b/a^2. \qquad (18)$$

where a, b are the coefficients in Bridgman's expression

$$-\Delta V/V_0 = aP - bP^2, \quad T \text{ constant}.$$

The coefficient b represents some average measure of the anharmonicity of the solid. But by using directly the results of Lazarus the adiabatic modulus can be inserted into the theory; taking $s \propto (\chi_s \rho)^{-1/2}$ gives

$$\gamma = -\frac{1}{6} + \frac{1}{2}\left(\frac{\partial \ln \chi_s}{\partial \ln V}\right)_T = -\frac{1}{6} + \frac{1}{2}\left(\frac{c_{11}{}^{s'} + 2c_{12}{}^{s'}}{c_{11}{}^{s} + 2c_{12}{}^{s}}\right). \qquad (19)$$

In table 3, γ is computed from eqns. (18) and (19) using Lazarus' data including the values of a and b he records for his five materials.

Table 3

Material	$-\frac{2}{3} + \frac{b}{a^2}$	$-\frac{1}{6} + \frac{1}{2}\left(\frac{\partial \ln \chi_s}{\partial \ln V}\right)_T$
Cu	1·83	2·22
Al	1·86	2·23
KCl	1·58	2·71
NaCl	2·00	3·07
CuZn	2·13	2·25

§ 5. Discussion of Results

When arranged according to increasing experimental γ's (table 4) it can be seen that the theoretical γ_L, γ_H also exhibit a general increase except for the case of Cu. As expected the actual numerical agreement is much better for γ_H than for γ_L, and for the alkali halides is well within the limits imposed by experimental error in the original data. But for Cu and Al some other explanation of the numerical discrepancy must be sought.

Table 4

	KCl	NaCl	Cu	CuZn	Al
γ_{expt}	1·48	1·60	2·03	2·22	2·26
γ_L	0·525	1·23	0·76	1·74	3·77
γ_H	1·57	1·60	0·937	1·76	3·26

It appears from table 2 that it is the contributions from the transverse branches which cause the final values of γ_H (and γ_L) to differ from the observed γ by so great an amount. Yet this does not offer an explanation of the discrepancies because the different values of γ associated with the polarization branches are still unexplained. Indeed, the continuum model does not attempt to explain them; it simply takes these values from direct observation of the low frequency end of the lattice spectrum.

Assuming the basic theory leading to eqn. (1) is sound, there are two possible causes of the discrepancies: errors in the data or inadequacies of the model. The elastic moduli are themselves accurate to within a few per cent but errors of 10 or 20% may reasonably be allowed in the third-order constants. Granting this, a discrepancy of a factor of 2 for the case of Cu still cannot be explained. The more likely explanation lies in the basic assumption of our model which allows us to replace the $\gamma_{\mathbf{q},p}$'s by the values at the long wavelength end of the spectrum. It would hardly be surprising if, in fact $\gamma_{\mathbf{q},p}$ were very different for short waves thus causing the average γ to differ markedly from that predicted by the continuum model.

The humiliating fact is that Slater's calculation, based upon the crudest possible model, gives better overall agreement with experiment, especially when one considers the assumptions in its derivation. Indeed, if one of his false assumptions is corrected, giving the modified Slater values of eqn. (19), the agreement is made worse. It is not difficult to see why these modified values differ from γ_H: they depend on the anharmonicity only through the volume variation of the adiabatic compressibility; it is effectively the average value of γ for compressive (longitudinal) waves that we are calculating as may be seen by comparing with the column γ_I in table 2.

From our results some tentative predictions as to the temperature variation of γ can be made. The Grüneisen parameter is a function of temperature (at say, constant pressure, to conform with experimental

measurements), this dependence being due to two distinct causes. The $\gamma_{\mathbf{q},p}$'s themselves may depend on temperature, for the elastic moduli and third-order constants are temperature dependent, but this effect is likely to be small. Of greater importance is the temperature dependence of the weighting factors $C_{\mathbf{q},p}$ of eqn. (1). This causes the weighted average of the $\gamma_{\mathbf{q},p}$'s to vary: at low temperatures, owing to the $1/s^3$ factor, γ is close to the value for the transverse branches, while at high temperatures the unweighted average is the appropriate one. If this is the main cause of the temperature variation the low temperature limit of γ should be close to our γ_L. Thus, for KCl and NaCl γ should decrease with decreasing temperature, the behaviour being more marked for KCl. Unfortunately, experiments are only available for the metals Cu and Al for which a slight decrease has been observed below 40°K (Bijl and Pullan 1955, Simmons and Balluffi 1957): our values of γ_L are too unreliable for any comparison to be made here.

Finally we remark that the electron gas in a metal only contributes to the Grüneisen parameter at temperatures low enough for the electronic and lattice specific heats to be comparable (Bijl and Pullan 1955): eqn. (1) need therefore only be corrected for temperatures below a few degrees Kelvin and the influence upon the room temperature γ's, with which this paper is concerned, is negligible.

Acknowledgments

I wish to express my thanks to Dr. J. M. Ziman, who suggested this calculation, for his encouragement and advice; and to Mr. M. F. Mitchell for his assistance in the numerical work. I am also indebted to the Department of Scientific and Industrial Research for the award of a maintenance allowance.

Note added in proof.—Since this paper was written information on the pressure variation of the elastic constants of Cu, Ag and Au has been published (Daniels and Smith 1958). For Cu the pressure derivative of the shear modulus, c_{44}^s, is larger by a factor of 3 than the earlier

Material	γ_H	γ_{expt}
Cu	1·97	2·03
Ag	2·40	2·48
Au	3·02	2·90

measurement and considerably alters the calculated Grüneisen parameter. In fact the values of γ_H derived from these data are in good agreement with the experimental values for all three metals indicating that the continuum model is of wider applicability than suggested by Lazarus' measurements.

References

Barron, T. H. K., 1955, *Phil. Mag.*, **46,** 720.
Berg, W. T., and Morrison, J. A., 1957, *Proc. roy. Soc.* A, **242,** 467.
Bijl, D., and Pullan, H., 1955, *Physica*, **21,** 285.
Birch, F., 1947, *Phys. Rev.*, **71,** 809.
Blackman, M., 1955, *Handbuch der Physik*, **7, 1.**
Buffington, R. M., and Lattimer, W. M., 1926, *J. Amer. chem. Soc.*, **48,** 2305.
Clusius, K., Goldman, J., and Perlick, A., 1949, *Z. Naturf.*, **4a,** 424.
Daniels, W. B., and Smith, C. S., 1958, *Phys. Rev.*, **111,** 713.
Dayal, B., 1944, *Proc. Indian Acad. Sci.*, A, **20,** 145.
Giauque, W. F., and Meads, P. F., 1941, *J. Amer. chem. Soc.*, **63,** 1897.
Grüneisen, E., 1926, *Handbuch der Physik*, **10, 1.**
Hearmon, R. F. S., 1953, *Acta cryst.*, **6,** 331.
International Critical Tables, **2,** 470.
Klemens, P. G., 1956, *Handbuch der Physik*, **14, 1.**
de Launay, J., 1956, *Solid State Physics* (New York: Academic Press), p. 266.
Lazarus, D., 1949, *Phys. Rev.*, **76,** 545.
Leibfried, G., and Schlömann, E., 1954, *Nachr. Akad. Wiss. Göttingen*, IIa, No. 4. 71.
Quimby, S. L., and Sutton, P. M., 1953, *Phys. Rev.*, **91,** 1122.
Simmons, R. O., and Balluffi, R. W., 1957, *Phys. Rev.*, **108,** 278.
Slater, J. C., 1939, *Introduction to Chemical Physics* (New York: McGraw-Hill), pp. 219, 238.
Sykes, C., and Wilkinson, H., 1937, *J. Inst. Metals*, **61,** 223.

Thermo-Remanent Magnetization (TRM) of Multidomain Grains in Igneous Rocks†

By F. D. Stacey
Department of Geophysics, Australian National University, Canberra, Australia

[Received July 23, 1958]

Abstract

The magnetic moment of a multidomain grain at its blocking temperature is determined by the domain arrangement which gives the minimum magnetic energy; for a large grain the energy minimum is produced by a self-demagnetizing field equal and opposite to the external field. Below the blocking temperature the domain configuration is frozen in and the moment increases with the increasing spontaneous magnetization. This results in a TRM proportional to small inducing fields, as in the theory of single domain grains and as required for agreement with experimental observations. Artificial TRM's of 17 rocks for which analyses of magnetic minerals are available show good agreement with calculated values. The theory also gives a formula for the maximum possible initial susceptibility of rocks which agrees well with experimental observations.

§ 1. Introduction

In the process of cooling of igneous rocks, grains of magnetic mineral are separated, generally as magnetite or titanomagnetite. The ferro-magnetic character of these grains appears at their Curie points, several hundred degrees below the solidification temperatures of the rocks, but although spontaneous magnetization occurs at these temperatures, as is evident from saturation magnetization measurements, the property of remanence or ability to retain induced moments is not acquired until temperatures several tens of degrees lower still are reached (Nagata 1953).

Thermal agitation is evidently responsible for preventing remanence above these lower temperatures, which have been termed 'blocking temperatures' (Néel 1955). The moment induced in a grain at its blocking temperature is 'frozen in' by further cooling and becomes thermo-remanent magnetization (TRM). This is important in rock magnetism because of its stability or resistance to demagnetization.

The general properties of TRM, including its proportionality to the inducing field, and the occurrence of blocking temperatures are well explained by the theory of Néel (1955), in which single domain grains of magnetic mineral are considered. However, Néel recognized that most naturally occurring magnetic grains were much too large to be correctly considered as single domains. The purpose of the present paper is to present a simple theory of TRM in multidomain grains, which has a direct and quantitative application to the magnetism of igneous rocks.

† Communicated by the Author.

For Néel's (1955) mechanism of thermal agitation of single domain grains to restrict the natural moments of igneous rocks to small fractions of their saturation moments we require

$$\frac{\mu H}{kT_B} \ll 1 \qquad \ldots\ldots\ldots \quad (1)$$

where H is the external field imposed on a grain which has a moment μ at its blocking temperature T_B and k is Boltzmann's constant. Except in the special case of a grain for which T_B is so close to the Curie temperature that the spontaneous magnetization has only a very small fraction of its value at ordinary temperatures, formula (1) can apply only to grains of the order 0·01 micron in diameter, whereas grains from one to 200 microns appear to be most important in the magnetism of igneous rocks (Nagata 1953). If such large grains were single domains they would acquire saturated TRM in very small fields, and the fact that this is not observed is particular evidence that each grain contains a number of domains.

§ 2. TRM in Large Grains

The total magnetic energy of a grain may be divided into the following six parts (Bozorth 1951, p. 811):

(i) exchange energy ;
(ii) crystal anisotropy energy ;
(iii) magnetic strain energy ;
(iv) domain wall energy ;
(v) magnetostatic energy (self-demagnetization due to magnetic poles) ;
(vi) mutual energy between magnetic moment of grain and external field.

We are interested in small fields, in which the moment is well below saturation and the domains are magnetized in easy directions, all of which have equal anisotropy energy in the absence of strain (Bozorth 1951, p. 478), so that (i) and (ii) are independent of the state of magnetization of the grain and can be neglected for the present purpose. In a large grain in its unmagnetized state the domains are so arranged that there are virtually no surfaces of magnetic polarity, either on the surface of the grain or internally (Bozorth 1951, p. 834), so that (v) is minimized. There is a further requirement that (iii) plus (iv) must have the minimum value, but if the grain is much larger than the limiting size for single domains, (v) exercises a much stronger control on domain geometry than (iii) and (iv) because the volumes occupied by closure domains and domain walls are small fractions of the whole grain. Therefore only (v) and (vi) need be considered for a first approximation to the domain structure of the grain.

In an external field, H_E, surfaces of magnetic polarity are induced on the surface of the grain, but not internally. This is equivalent to giving

the grain a moment I per unit volume, and results in a self-demagnetizing field $(-NI)$, N being the demagnetizing factor, a geometrical quantity explicitly stated only for ellipsoidal grains. In this state the grain experiences an internal field H_I, given by:

$$H_I = H_E - NI. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Contributions (v) and (vi) to the magnetic energy give a total value

$$E = -H_E I + \tfrac{1}{2} N I^2, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where the first term is (vi) and the second term is (v). If the grain is above its blocking temperature, the domain readjustments in the field H_E are not subject to internal restraints, so that I follows H_E reversibly and is self-determined to give E the minimum value. This is found by putting

$$\frac{dE}{dI} = -H_E + NI = -H_I = 0,$$

whence

$$I = \frac{H_E}{N}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

That this corresponds to a minimum of E is shown by a second differentiation

$$\frac{d^2E}{dI^2} = N,$$

which is necessarily positive.

As an alternative derivation of eqn. (4) we may assume that, for small external fields, the permeability of the material of the grain is effectively infinite above its blocking temperature, which is equivalent to the assumption that there are no restraints to domain readjustment, and implies directly that the internal field H_I is zero.

As long as H_E is small, so that I is well below saturation, the equilibrium described by eqn. (4) is maintained during the cooling of the grain until the blocking temperature is reached and the relaxation time for the decay of moments not in equilibrium with H_E becomes long compared with the time scale of cooling. Below this temperature the moment increases in the same way as the spontaneous magnetization, σ, so that at laboratory temperatures the thermo-remanent moment is

$$I_{TRM} = \frac{H_E}{N} \cdot \frac{\sigma_R}{\sigma_B}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where (σ_R/σ_B) represents the ratio of spontaneous magnetizations at room temperature and the blocking temperature.

As an approximation to the TRM of the magnetic minerals in rocks, eqn. (5) gives a reasonably good fit to experimental results, but detailed comparison will be left until a refinement has been added in § 4.

§ 3. Susceptibility of Rocks

The susceptibility of rocks at laboratory temperatures will now be considered. This requires a re-examination of the term 'blocking

temperature' which has been adopted from Néel's (1955) theory of TRM in single domain grains.

The moment of a single domain grain can change direction under the influence of thermal agitation, but only in discrete jumps because it is constrained to lie in one of a few directions by shape and crystal anisotropies. The blocking temperature is the temperature at which thermal agitation provides just sufficient energy for the moment to cross the potential barriers between these directions. Below this temperature the probability that a change in domain direction will occur becomes so small that in effect the moment is completely frozen in. Therefore a single domain grain below its blocking temperature cannot contribute to the low field susceptibility of a rock.

The susceptibility of a multidomain grain at low temperatures is not zero because a small moment may be induced in the grain by a continuous and reversible process of domain boundary displacements and not in large, discontinuous jumps. This is the process used to explain the initial susceptibility of large ferromagnetic bodies (Bozorth 1951, p. 480). Its reversibility arises from internal restoring forces which are due to magnetic energy contributions (iii) and (iv) (§ 2). These are constraints to domain readjustment which prevent the material of the grain from having infinite susceptibility. At high temperatures thermal agitation overcomes these constraints and allows the domain geometry to readjust itself to give the minimum total magnetic energy. The blocking temperature of such a grain may then be defined as the temperature at which internal constraints become effective in limiting the adjustment of domain boundaries when the grain is subjected to small fields.

Accurate estimates of the susceptibilities of rocks cannot be made because the internal forces limiting domain readjustment are not well understood. However, the maximum possible susceptibility of a rock may be estimated by the following argument. The condition for this maximum to be observed is that internal constraints have only a negligible effect in limiting domain wall movement, and the moment induced in a grain has that value which results in a minimum of the total magnetic energy, being given by eqn. (4). Putting $N = (\frac{1}{3}\pi)$ and considering a rock with M% by weight of magnetic minerals of density 5·25 g/cm^3, we have for the moment per cm^3 of the rock:

$$\frac{I}{H_E} = \frac{3}{4\pi} \cdot \frac{M}{100} \cdot \frac{\delta}{5{\cdot}25},$$

where δ is its density. The maximum possible mass susceptibility of the rock is then given by

$$\max \chi_M = \frac{I}{\delta H_E} = 0{\cdot}45 \times 10^{-3} M. \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Nagata (1953, p. 93) showed that the maximum initial susceptibility of igneous rocks is given approximately by

$$\max \chi_0 = 4{\cdot}5 \times 10^{-2} C_{Mt},$$

where Nagata's χ_0 is the same as χ_M in the present paper and C_{Mt} is the fraction by weight of normative magnetite in a rock. The agreement with eqn. (6) is remarkably good if M is the percentage of normative magnetite and not the total iron–titanium oxide content.

More useful is a comparison of eqn. (6) with the maximum susceptibilities of rocks observed at elevated temperatures, at which the internal constraints to domain wall movement are least effective. This comparison is made in table 1, using data of Nagata (1953) taken from his table 3–IV. The calculated values of maximum susceptibility have been obtained, using normative magnetite contents for values of M in eqn. (6). It is not clear why this gives better agreement than the total Fe–Ti oxide content which indicates a maximum initial mass susceptibility of the order 6×10^3 for the rocks listed.

Table 1. Calculated and Observed Maximum Susceptibilities of some Japanese Rocks. Specimen Numbers and Experimental Values are from Nagata (1953)

Specimen number	(max χ_M) calc. $\times 10^3$	(max χ_M) obs. $\times 10^3$
17	3·54	2·56
18	1·77	2·06
19	2·40	2·75
20	2·92	1·42
21	3·13	1·93
22	1·77	1·08
23	1·98	1·46
24	0·62	0·68
25	1·15	2·05
28	1·15	1·42
30	1·98	1·60
32	1·77	0·79
33	0·83	0·15
34	1·04	0·96
35	1·35	0·42
36	2·71	2·48
37	3·02	3·34
38	1·76	2·06
51	1·67	0·84
56	2·29	0·87
91	1·67	2·02

§ 4. Reduction of TRM by the Self-Demagnetizing Field

The self-demagnetizing field exerted on a grain by its moment I_0 per unit volume is $(-NI_0)$, so that the value of I_{TRM} calculated in § 2 is reduced to I_0 by this field, where

$$I_{TRM} - I_0 = NI_0\chi_V,$$

χ_V being the volume susceptibility of the material of the grain at laboratory temperatures. The observed TRM, I_0, is thus given by

$$I_0 = \frac{I_{TRM}}{1+N\chi_V} . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

Now we can estimate χ_V for the material of the grains in a rock from the observed susceptibility of the whole rock. The moment I induced in a grain by an external field H_E is given by

$$I = \chi_V(H_E - NI) = \frac{\chi_V H_E}{1+N\chi_V} .$$

If χ_M is the mass susceptibility of the rock, density δ, with $M\%$ by weight of magnetic minerals of density δ_M then

$$\chi_M = \frac{I}{\delta H_E} = \frac{\chi_V}{1+N\chi_V} \cdot \frac{(0{\cdot}01M)}{\delta_M} ,$$

which gives

$$\chi_V = \frac{1}{(0{\cdot}01M/\chi_M\delta_M) - N} . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

Substituting (8) into (7) we obtain

$$I_0 = \frac{I_{TRM}}{1+[(0{\cdot}01M/\chi_M\delta_M)(1/N)-1]^{-1}} . \quad . \quad . \quad . \quad . \quad (9)$$

The value of I_{TRM} for the whole rock is obtained by multiplying eqn. (4) by $(0{\cdot}01M/\delta_M)$ so that eqn. (9) becomes:

$$I_0 = \frac{(H_E/N)(\sigma_R/\sigma_B)0{\cdot}01M(\delta/\delta_M)}{1+[(0{\cdot}01M/\chi_M\delta_M)(1/N)-1]^{-1}} ,$$

whence

$$\frac{I_0}{\delta H_E} = \frac{\sigma_R}{\sigma_B}\left(\frac{0{\cdot}01M}{\delta_M} \cdot \frac{1}{N} - \chi_M\right) . \quad . \quad . \quad . \quad . \quad (10)$$

Equation (10) expresses the TRM per gram of rock per oersted in terms of the weight percentage M of magnetic mineral of density δ_M, and χ_M is the mass susceptibility of the rock.

§ 5. Comparison with Experiment

Equation (10) gives the required proportionality of TRM and field and also allows a reasonably accurate estimate to be made of the constant of proportionality. Using data of Nagata (1953), the average blocking temperature of an assembly of grains is usually 60°C ± 20°C below the Curie temperature, and his fig. 1-41 indicates $(\sigma_B/\sigma_R) \sim 0{\cdot}3$ for a blocking temperature 60° below the Curie point of a mineral which appears to be nearly pure magnetite. This ratio will be subject to some variation but the value 0·3 is probably accurate within a factor of two for most igneous rocks. It is to be expected that the magnetic grains in igneous rocks have randomly oriented shape anisotropies, so that $N = (\frac{1}{3}\pi)$ as for spherical grains will be taken as an approximate effective value; this is

reconsidered in § 7. The value of δ_M is taken as 5·25 g/cm³, as for pure magnetite (Nagata 1953); the densities of other magnetic minerals do not depart very widely from that of magnetite. Substitution of these values into eqn. (10) reduces it to:

$$\frac{I_0}{\delta H_E} = 1{\cdot}5_{16} \times 10^{-3} M - 3{\cdot}3_3 \chi_M . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

Values of $(I_0/\delta H_E)$ calculated from eqn. (11) are compared with experimental values in table 2, using data of Nagata (1953, tables 3–I, 4–V, 4–VI), whose specimen numbers are given in the first column. $(I_0/\delta H_E)$ corresponds to I_{Tc} and χ_M to χ_0 in Nagata's tables.

Table 2. Calculated and Observed TRM's of some Japanese Rocks. Specimen Numbers and Experimental Values are from Nagata (1953)

Specimen number	% Magnetic Mineral†	$\chi_M \times 10^3$	$(I_0\ \delta H_E)$ calc. $\times 10^3$	$(I_0\ \delta H_E)$ obs. $\times 10^3$
17	13·45 (7·87)	2·25	12·9 (4·4)	12·1
18	11·89 (3·94)	1·31	13·7 (1·6)	10·7
19	15·54 (5·33)	1·82	16·5 (2·0)	9·7
20	13·17 (6·48)	1·28	15·7 (5·6)	10·5
21	11·88 (6·95)	1·55	12·8 (5·4)	13·0
22	12·44 (3·94)	0·70	16·5 (2·6)	41·7
23	11·48 (4·40)	1·29	13·1 (2·4)	5·2
24	6·84 (1·39)	0·40	9·0 (0·8)	7·9
25	11·60 (2·55)	0·90	14·6 (0·9)	14·8
27	11·19 (4·17)	1·28	12·7 (2·1)	3·7
38	14·00 (3·91)	1·01	17·9 (2·6)	27·6
52	10·76 (4·63)	1·09	12·7 (3·4)	2·8
53	8·78 (3·94)	0·57	11·4 (4·1)	9·0
93	10·22 (3·47)	0·54	13·7 (3·5)	4·6
94	10·64 (6·02)	0·78	13·5 (6·5)	4·8
95	14·91 (6·25)	1·37	18·0 (4·9)	5·1
96	12·24 (5·09)	0·73	16·1 (5·3)	4·6

† Figures are total Fe–Ti oxide content by weight; values obtained using normative magnetite are given in parentheses.

It is not immediately obvious whether the total iron–titanium oxide content or the normative magnetite gives the more appropriate value of M. If the whole Fe–Ti content occurs in strongly magnetic grains then the former is correct. However, in many rocks some of the Fe–Ti content occurs in non-magnetic minerals which do not contribute to the TRM, and for such rocks the normative magnetite content may give a better value of M. Columns 4 and 5 of table 2 indicate that in ten of the rocks listed the whole Fe–Ti content is magnetically active, but that in the other seven a separation of magnetite has occurred.

§ 6. Decay of TRM with Time

A matter of some geophysical interest arises from the theory which has been presented. The evident decay over geological time of the natural remanent moments of rocks has been regarded (Nagata 1953, p. 140) as a decay of natural TRM, which at ordinary temperatures has a very long relaxation time. It is imagined that thermal agitation causes a gradual disruption of TRM. If this were the case then it would be exactly the same process as that by which TRM is acquired, except that it takes a very much longer time, and therefore remanent moments with the stability of TRM would be gradually induced by the changing geomagnetic field at the same time as the original moments decayed. Rocks with decayed natural moments would then be quite unsuitable for palaeomagnetism.

However, the TRM of multidomain grains causes self-demagnetizing fields within the grains which are several times stronger than the field in which the TRM is induced. These self-demagnetizing fields must be the principle cause of decay of TRM. They are exactly opposite to the moments of the grains and the decay of the moments therefore involves no changes in their directions. Since the influence of a changing geomagnetic field is only secondary, rocks with decayed natural moments are not necessarily unreliable for palaeomagnetic work.

§ 7. Discussion

Further consideration will be given to the following approximations or assumptions which have been made:

(1) multidomain grains have blocking temperatures;

(2) the susceptibility of the material of a grain above its blocking temperature is so high that the self-demagnetizing field provides the only limit to the magnetic moment;

(3) an average demagnetizing factor ($\frac{4}{3}\pi$) can be used for all grains;

(4) grain interactions can be neglected;

(5) grains of magnetic mineral are much larger than the limiting size for single domains;

(6) the ratio (σ_R/σ_B) is roughly constant for all rocks;

(7) chemical analysis indicates the content of magnetic minerals.

(1) A definition of the term 'blocking temperature' as applied to multidomain grains was suggested in § 3, but it requires qualification, since it is not a single temperature but a temperature range. The crystal imperfections and internal strains which become effective in opposing changes in domain geometry at low temperatures are unlikely to be distributed so uniformly through a grain that movements of all domains are opposed equally. It is more probable that some domain boundaries become frozen in at higher temperatures than others within the same grain; the maximum spread can be inferred from the sharpness of the blocking

temperature of a sample of artificial magnetite powder (Akimoto, quoted by Nagata 1953, p. 161) of uniform grain size, and appears to be of the order 100°c. This is not inconsistent with an average blocking temperature 60°c below the Curie point as assumed in § 5, so that the estimated value of (σ_R/σ_B) is probably not seriously in error on this account.

(2) This assumption has been discussed by Néel (1955) who concluded that "the apparent susceptibility of rocks with coercive forces of less than 500 oersteds... must depend more on the demagnetizing field effects than on the intrinsic susceptibility of the material"; coercive forces are certainly less than 500 oersteds at elevated temperatures.

(3) The value $(\frac{4}{3}\pi)$ for the average demagnetizing factor in the field direction of a random assembly of grains is only an approximation and in general a somewhat smaller value must be expected. This can be seen by considering two very elongated grains whose major axes make equal, but opposite angles, θ, to the external field, H. Each grain acquires a TRM almost at an angle θ and given by

$$I_{\mathrm{TRM}} = \frac{H\cos\theta}{N'} \cdot \frac{\sigma_R}{\sigma_B},$$

where N' is the small demagnetizing factor of each grain. The net moment in the field direction due to the TRM of many such pairs of grains is then

$$I_{\mathrm{TRM}} = \frac{H}{N'} \cdot \frac{\sigma_R}{\sigma_B} \int_0^{\pi/2} \cos^2\theta \sin\theta \, d\theta,$$

$(\sin\theta)$ being the weighting factor for the number of grains at an angle θ. Thus

$$I_{\mathrm{TRM}} = \frac{1}{3}\frac{H}{N'}\frac{\sigma_R}{\sigma_B}$$

where N' may be much smaller than $(\frac{4}{3}\pi)$.

Generally dimension ratios greater than 3 or 4 are not to be expected; if all the grains in a rock are prolate ellipsoids with the dimension ratio 3·5, then the TRM parallel and perpendicular to the major axes can be calculated as above and the total TRM is 50% higher than that indicated by eqn. (11).

(4) In §§ 2–5 magnetic interactions between grains have been neglected. These interactions were considered by Néel who showed that under certain conditions they could lead to a mechanism of self-reversal of TRM. However, one of the conditions was a higher concentration of magnetic grains than is normally found in rocks. Assuming that the magnetic grains are uniformly dispersed in a rock, interactions between them cause a reduction in the effective self-demagnetizing fields of the individual grains. For spherical grains this reduction can be expressed in terms of an additional magnetizing field H' given by

$$H' = (\tfrac{4}{3}\pi) I',$$

where I' is the moment per unit volume of the rock. If I_0 is the moment which the same grains would give if they were very widely separated then

$$I' = I_0 + \delta \chi_M H',$$

δ and χ_M being the density and mass susceptibility of the rock. Thus

$$\frac{I'}{I_0} = \frac{1}{1 - (\frac{4}{3}\pi)\delta\chi_M}.$$

if $\chi_M \sim 10^{-3}$, as for the rocks listed in table 1, $(I'/I_0) \sim 1{\cdot}01$, so that unless the grains are arranged in a grossly non-uniform way this effect is negligible.

(5) It must be noted that the theory presented in §§ 2–5 applies only to large grains. Small grains, containing only a few domains, for which the absence of internal surfaces of magnetic polarity cannot necessarily be assumed or for which domain wall and magnetic strain energy cannot be neglected, must retain some of the behaviour of single domains, and therefore become magnetized to an extent greater than is indicated by eqn. (11). This probably applies to grains smaller than one micron in diameter.

(6) Assuming that the difference between the blocking temperature and the Curie point has been correctly estimated, the only variation in (σ_R/σ_B) arises from the differences between the shapes of spontaneous magnetization versus temperature curves. As an extreme case we may consider a grain with a blocking temperature of 520°C and spontaneous magnetization decreasing linearly between 20°C and the Curie point at 580°C. Such a grain would have $(\sigma_R/\sigma_B) \sim 9{\cdot}3$ compared with the value 1·9 estimated from the curve for nickel. These represent extremes and it is more reasonable to expect

$$2{\cdot}5 < \frac{\sigma_R}{\sigma_B} < 5.$$

(7) Magnetic analysis would be more appropriate than chemical analysis since both the magnetic minerals and their concentrations can be estimated from saturation magnetization versus temperature curves: also (σ_R/σ_B) could be found separately for each constituent. Nagata (1953, p. 107) has found that the saturation magnetization of rocks is roughly proportional to their normative magnetite, but that this does not represent the actual mineral content is evident from the constant of proportionality, which is only half the saturation magnetization of pure magnetite. Therefore the separation of magnetite referred to in § 5 must be a separation of titanomagnetite with saturation magnetization roughly half that of pure magnetite. The saturation magnetization data would be satisfied by a separation of pure magnetite to the extent of half the normative magnetite, but according to eqn. (11) this would give a TRM much smaller than is observed. Although no conclusion can be drawn, the results of table 2 are consistent with a separation of titanomagnetite having about 50 mol% $TiFe_2O_4$ particularly in specimens 27, 52, 93 to 96. Nagata (1953) noted that numbers 27, 52, 53 and 93 behaved irreversibly when heated, probably on account of a separation process, but that numbers 17 to 25 did not.

Acknowledgments

I wish to acknowledge the helpfulness of discussions with Professor J. C. Jaeger and Mr. E. Irving during the preparation of this paper.

References

Bozorth, R. M., 1951, *Ferromagnetism* (New York: D. Van Nostrand Co. Inc.).
Nagata, T., 1953, *Rock Magnetism* (Tokyo: Maruzen Co. Ltd.).
Néel, L., 1955, *Advanc. Phys.*, **4**, 191.

Adhesion of Evaporated Aluminium Films†

By C. Weaver and R. M. Hill
Department of Natural Philosophy,
The Royal College of Science and Technology, Glasgow

[Received August 2, 1958]

Abstract

The increase in adhesion obtained for aluminium films by the pre-deposition of films of chromium is investigated, using Heavens' method of adhesion testing. It is found that the adhesion varies with the thickness of the chromium film, and with time. The variation with time suggests that the nature of the increase in adhesion is primarily one of alloying at the metallic interface, and subsequent age-hardening.

§ 1. Introduction

Vacuum-deposited films of aluminium on glass are extensively used as front surface mirrors for optical purposes and it has been known for some years that the durability of these mirrors could be considerably increased by depositing a thin layer of chromium on the glass immediately before the aluminium (Holland 1956), both evaporations being performed in the same vacuum chamber without admitting air. The reflectivity of the aluminium film, which is relatively thick, is unaffected; but the adhesion to the glass is improved and it has been suggested that the chromium forms a barrier, preventing attack of the aluminium by the alkali in the glass (Holland 1956) with consequent loosening of the bonding between the aluminium. Such a screening effect would only prevent deterioration of the reflecting film and cannot be regarded as a complete explanation since an increase in adhesion is actually observed. A preliminary investigation into the nature of this increase in adhesion by Heavens (1950) showed that large increases in adhesion could be obtained when the thickness of the chromium deposit was greater than about 300 Å. It was later suggested (Heavens and Collins 1952) that this was due to an oriented overgrowth mechanism but it appears from the present work that the observed increases are largely due to a diffusion–precipitation phenomenon at the interface between the aluminium and the chromium.

It is difficult to measure the adhesion of thin films and almost all accepted methods fail when the adhesion is strong. The results of abrasion testing (Townsley 1945) are not reliable since a burnishing action is involved, which affects results. Consequently, in this work, the adhesion of the evaporated films has been measured using the method adopted by Heavens. A chrome-steel point having a smoothly rounded

† Communicated by the Authors.

tip was drawn across the film and the load on the point was gradually increased until a critical load was reached at which the film was completely removed from the glass substrate leaving a clear channel in the film. This stage was readily determined by viewing the scratches under a low power microscope using transmitted light. At the critical load a broad evenly illuminated channel with sharply defined edges could be observed whereas lower loads did not remove the metal completely and a scratch was produced which appeared bright at the centre, tapering to complete opacity at the edges, indicating a channel of more circular cross section. The results obtained by this method were readily repeatable and an accuracy of $\pm 5\%$ could be obtained in determining the end-point.

Heavens did not give any analysis of his method but the application of the method to single metal films on various substrates has since been investigated by Weaver and Benjamin (1958) who showed that it gives an effective measure of adhesion. They have shown that the shearing force at the interface is a direct function of the vertical load on the point and that calculated results are in good agreement with theoretical estimates. The validity of the method for double films such as aluminium upon chromium has not, however, been established and will be discussed later in the light of experimental results.

§ 2. Experimental

Specimens were prepared on glass microscope slides by evaporating chromium and aluminium in succession so that each metal extended over two-thirds of the length of the slide, overlapping in the centre. Thus the centre section of the slide was covered first with a chromium film and then by an aluminium film while the two end sections were covered with pure chromium and pure aluminium respectively. The evaporations were performed in a glass bell-jar evacuated by an oil diffusion pump, which produced pressures of $2{\cdot}5 \times 10^{-5}$ mm Hg as measured by an ionization gauge. The slides were supported in a rotating jig, about 20 cm above the heaters which were made of tungsten wire, 1 mm diameter, wound in the form of spirals. The slides were cleaned carefully before placing them in the chamber, being scrubbed first with Teepol solution to remove dirt and grease and then rinsed under running water before drying on a linen cloth and finishing with lens tissue. One position of the rotating jig allowed the full surface of the slide to be exposed to a cleaning discharge in the bell-jar for about 15 minutes during the pumping process. When a sufficiently low pressure had been reached the jig was rotated so as to bring the slide first over the chromium heater and later, over the aluminium heater. Suitable masks were used so that only two-thirds of the slide near one end were coated at a time. The rates of evaporation were approximately 10 Å/sec for the chromium and 20 Å/sec for the aluminium, the two evaporations being performed in rapid succession.

The thicknesses of the chromium films were monitored during deposition and measured *in vacuo* by using a parallel beam of white light and

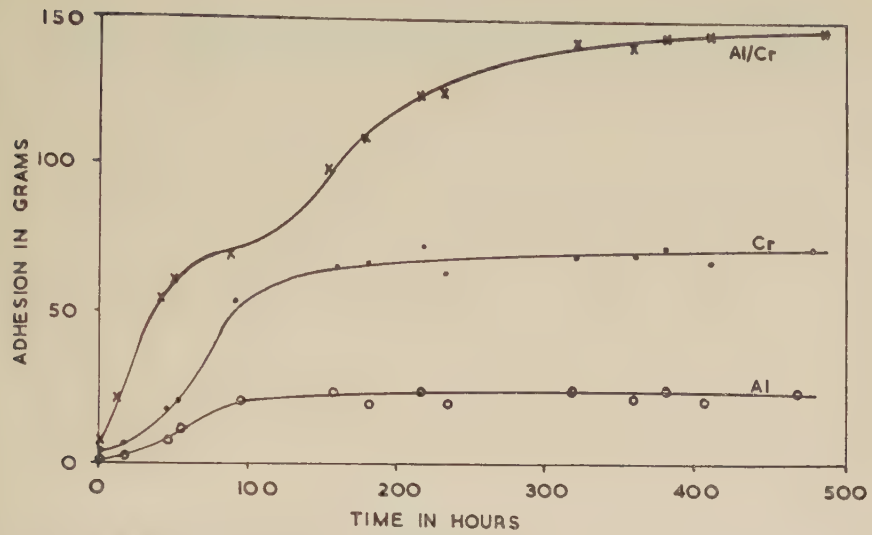
measuring the transmission. The light beam was projected into the vacuum chamber and reflected by a glass prism, screened from the evaporated material, so that it passed through the specimen slide and struck a barrier layer photocell connected to a sensitive galvanometer. A calibration curve of film thickness against transmission was obtained by weighing a number of specimen slides, covered with chromium only, on a microbalance immediately after deposition. The thickness was calculated from the bulk density so that the measurements gave an effective weight thickness. The aluminium films were evaporated to be just opaque, corresponding to a thickness of about 600–800 Å.

Immediately after preparation of each specimen it was taken from the vacuum chamber and measurements of adhesion were made. Values were obtained for aluminium, chromium and for the double film of chromium overlaid by aluminium on the same slide and prepared under the same conditions. From the relative values the effect of the chromium underlayer could be seen. It was known that commercially produced films were aged after deposition and it was noticed during the investigation that the adhesion values obtained changed with time. Measurements were therefore made periodically until steady readings were obtained and ageing was complete. For films kept at room temperature this ageing process took about three weeks. In accordance with commercial practice some of the specimens were aged in ovens at a temperature of about 120°c and, in this case, the ageing process was completed in about three days.

§ 3. Results

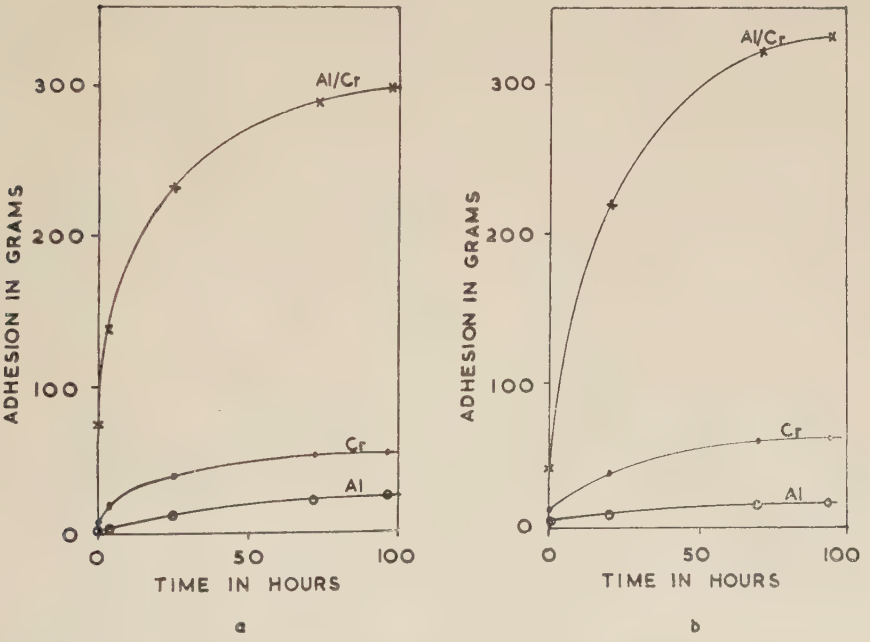
Typical ageing curves obtained for one specimen which had been aged at room temperature are shown in fig. 1. The initial adhesion figures obtained both for the pure metals and for the double film were all low. The adhesion of the aluminium film increased gradually with time but quickly reached a steady but low value. The final value obtained did not vary noticeably with thickness or from one specimen to another. The adhesion of the pure chromium also increased steadily but more rapidly at first and over a somewhat longer time, finally reaching a steady value. This value varied slightly with the thickness of the specimen, being somewhat greater for thicker films. The significant feature of these ageing curves is, however, the difference between the chromium curve and the curve obtained for the double film of aluminium overlying chromium. Chromium undoubtedly sticks to glass better than aluminium and this could partially explain the increased adhesion of the double film, as compared with aluminium, but the difference between the two ageing curves must still be explained. The difference is even more pronounced for specimens aged at 120°c and typical curves are shown in figs. 2 (*a*) and 2 (*b*). The final steady values obtained for pure aluminium and pure chromium are comparable with the values obtained when ageing at room temperature. The final values reached by the double films are, however, much greater and the shape of the ageing curve has changed.

Fig. 1



Ageing of a film of aluminium–chromium at room temperature. Adhesion–time curves. Thickness of aluminium 700 Å; thickness of chromium 150 Å.

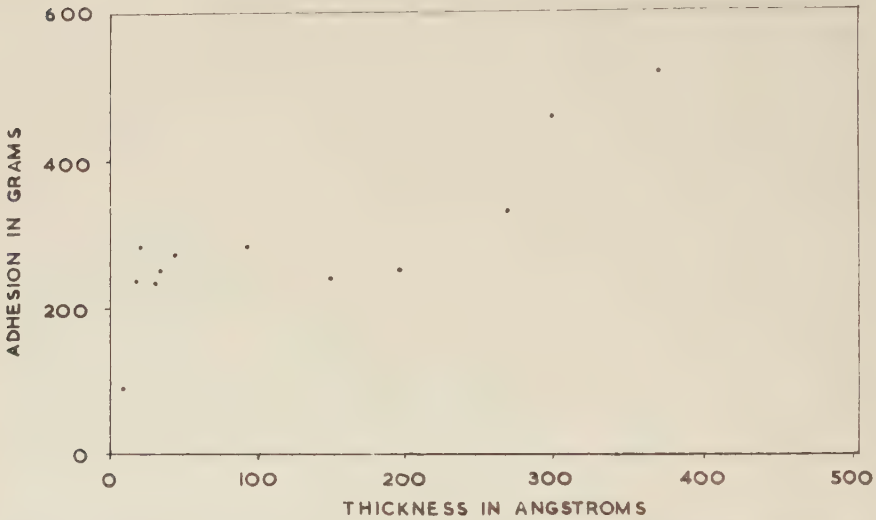
Fig. 2



Ageing of a film of aluminium–chromium at 120°c. Adhesion–time curves.
(*a*) Thickness of aluminium 600 Å; thickness of chromium 50 Å.
(*b*) Thickness of aluminium 600 Å; thickness of chromium 250 Å.

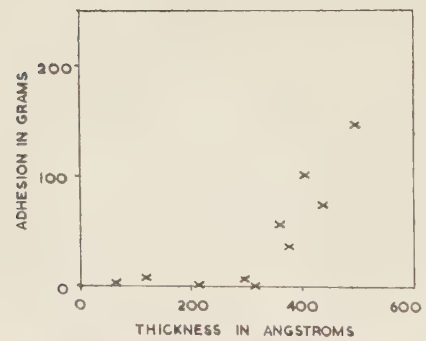
The final steady value of the adhesion of a double film after ageing varies considerably with the thickness of the chromium underlayer; this is shown in fig. 3, where results obtained from a large number of specimens are given. This graph does not compare very well with the results

Fig. 3



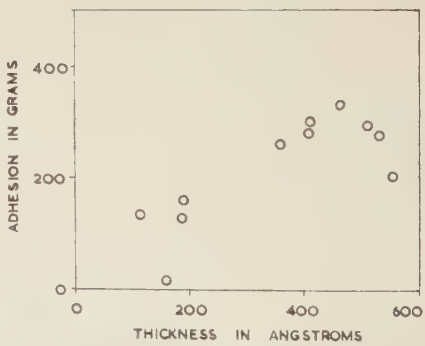
Graph of adhesion–thickness of underlayer for films of aluminium -chromium, aged at 120°C.

Fig. 4



Graph of adhesion–thickness obtained by Heavens.

Fig. 5



Graph of adhesion–thickness obtained immediately after preparation.

obtained by Heavens, as shown in fig. 4. His results compare more favourably with fig. 5 in which we have plotted the adhesion values obtained immediately after evaporation, against the thickness of the chromium underlayer. This suggests that Heavens did not take ageing effects into account but measured his specimens immediately after removal from the vacuum chamber.

§ 4. Discussion

The increased strength of the double metal film cannot be explained in terms of any change in the basic adhesion forces between the chromium and the glass. The same chromium films act as a single metal film over one section of the slide and as an underlayer for another section. This implies that the difference between the measurements on the two sections is due to the aluminium overlayer and for the aluminium to affect the interface between the chromium and the glass, aluminium atoms would have to diffuse through the chromium film. This would be a slow process compared with the rapid initial rise of the ageing curves, even if we assumed that the aluminium atoms at the interface could increase the adhesion. This seems most unlikely since our measurements show that chromium adheres to glass better than aluminium.

The ageing process for the double films at room temperature normally appears to consist of two stages occurring at different times. Similar processes have been observed in the age hardening of alloys and our curves correspond closely with those of Gayler (1937) who attributed the two stages of ageing to an interdiffusion of the two metals followed by a partial precipitation of an intermetallic compound. According to the theories of age hardening (Smith 1949), the interdiffusion would cause a hardening or increased resistance to deformation with no change in lattice parameters and this would be followed by a release of strain and precipitation of an intermetallic compound. At higher temperatures the processes would be accelerated and could overlap, giving one continuous change such as is observed in our graphs of high temperature ageing.

There appears to be some doubt about the phase diagram of the aluminium–chromium system (Metals Handbook 1948), and information is not complete but it has been claimed that there are nine distinct intermetallic compounds (Bradley and Lu 1937). Chromium does, however, have strong similarities to nickel and cobalt and complete phase diagrams are available for aluminium–nickel and aluminium–cobalt. In each case the phase diagram has a pronounced peak indicating formation of a stable intermetallic compound of high melting-point. These compounds have a high energy of formation and are known to be hard and strong. It would therefore not be unreasonable to expect that at least one of the chromium–aluminium compounds will have similar properties. This leads us to suggest that, when a film of aluminium is deposited on top of a film of chromium, an intermetallic reaction occurs at the interface and a compound precipitates forming an intermediate layer. Such a process could explain the results which we have obtained since precipitation of a hard intermetallic compound would explain the second step in the room temperature ageing curves. In the method of adhesion measurement which has been used the effect of a strong intermediate layer between the aluminium and the chromium would be to spread the concentrated load upon the point over a wider area so that a greater force would be required to produce the same shearing effect at

the interface between the glass and the chromium. At the same time the bonding between the aluminium and the chromium would be improved. This implies, of course, that although this method of measuring adhesion has been shown to be suitable for single metal films and capable of theoretical interpretation, this is no longer true for double films such as are considered here. Valuable information may nevertheless be drawn from the results provided care is taken in the interpretation.

§ 5. Electron Diffraction Investigation

An electron diffraction investigation was carried out on some specimen films. These were prepared in the same apparatus as the adhesion slides, but the metals were condensed on to collodion films supported on copper grids. The table of the apparatus was designed to give films of the three types already investigated, i.e. both the pure metals on separate grids and a composite film on a third. An electron microscope was used to obtain the transmission electron diffraction patterns. Before the films were examined they were allowed to age for about 300 hr at room temperature.

The first set of grids examined consisted of a chromium film approximately 300 Å thick, and an aluminium film 500 Å thick. Both of the single metal films showed the characteristic diffraction pattern of the bulk materials, fig. 6 (*a*, *b*), Pl. 98. The compound film, fig. 6 (*c*), Pl. 98, gave a pattern that was simply a summation of those obtained for the single metals. No trace of intermetallic compounding was detected. If, however, the diffusion rates were low, and the degree of penetration small then diffraction effects from such a layer might be obscured by the relatively thicker metallic layers.

Another set of films was thus prepared using much thinner films, the thickness of the chromium layer being 30 Å, and of the aluminium 50–75 Å. After the ageing period, the single aluminium film showed only the presence of an amorphous substance, fig. 7 (*a*), Pl. 98, probably aluminium oxide. No trace of the metal itself was obtained. This is in agreement with the theories of oxide growth, since it has been estimated that the thickness of an oxide layer on aluminium would achieve the value of approximately 35 Å in 300 hours at room temperature (Mott 1947). The chromium pattern, fig. 7 (*b*), Pl. 98, also showed the presence of oxidation, but some of the metal has remained unoxidized.

For this specimen, the compound metal film gave a pattern completely uncharacteristic of any of the previous specimens, fig. 7 (*c*), Pl. 98. The rings are broad but distinct, representing a face-centred lattice of unit cube cell and spacing 2·92 Å. This corresponds to the α phase of the aluminium–chromium alloy system (Bradley and Lu 1937) and appears for a concentration of approximately 87% of chromium by atomic volume. This is not the most prominent phase in the system, but when it is considered that the structures of the other phases are based upon the stacking of three of the α phase unit cells on top of each other, and due to the direction of interdiffusion that the electron beam would more

likely traverse this system parallel to the '*c*' axis, then we can see that it is doubtful whether more information could have been obtained.

§ 6. Conclusion

The evidence deduced by electron diffraction appears to give reasonable confirmation of the general validity of our interpretation of the ageing curves which have been obtained for double films of aluminium on chromium. Such values as we have measured for the adhesion of these films immediately after deposition are in general agreement with the results given by Heavens in that films of chromium thicker than about 300 Å cause a marked increase in the adhesion of aluminium films. But even thinner chromium films cause some increase in adhesion according to our measurements although not so great as that produced by thicker chromium films. Heavens found that thin chromium films did not have any appreciable effect and we can only suggest that this may be due to differences in cleaning methods or evaporation conditions whereby our chromium films had a greater basic adhesion than those prepared by Heavens.

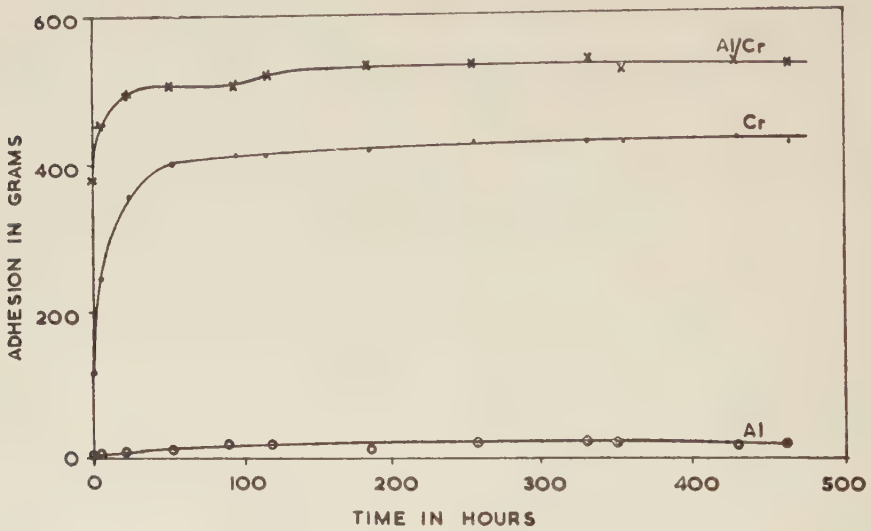
Heavens also indicated that in a case where air was admitted to the vacuum chamber after evaporating the chromium and before evacuating again to deposit the aluminium, no increase in adhesion was obtained. Such a result is to be expected if our theory is correct, since an oxide film would form rapidly on the chromium when exposed to air and such an oxide film would at least hinder if not prevent an intermetallic reaction such as we have suggested.

The existence of a critical thickness for these chromium underlayers, cannot readily be explained from the results. Previous work on the optical properties of thin films of chromium (Hill and Weaver 1958) has, however, shown that thin films of chromium are composed of aggregates which may be regarded as oblate spheroids with the minor axis perpendicular to the substrate. Each aggregate is surrounded by a very thin layer of oxide when examined in air, and the size of the aggregates changes with increasing film thickness, the axial ratio becoming greater so that the aggregates tend to become platelets. This type of structure exists up to weight thicknesses of about 100–120 Å, above which there is a transition region and the structure changes to a more continuous type corresponding to metallic chromium. It appears quite feasible that films below about 200–300 Å have not reached the stage of having a continuous metallic structure even in their upper layers, whereas it has been shown that the upper layers of films of about 500 Å weight thickness have the same refractive index as bulk chromium metal. In thin films the screening of the aggregates by a layer of oxide would prevent an intermetallic reaction with a superimposed aluminium film such as has been suggested and, consequently, these films would have little effect in increasing adhesion.

Finally, we ought to point out that since obtaining the results given it has been found that the basic adhesion of the chromium films themselves

can be increased by using a shutter in front of the source and moving it only when the source is at operating temperature so that only high velocity chromium particles strike the substrate to form a deposit. Under these conditions the overall adhesion of a double film is increased and the effect of the intermetallic reaction becomes less noticeable but it nevertheless still persists and plays some part as can be seen from the ageing curves of fig. 8.

Fig. 8



Ageing of a film of aluminium–chromium at room temperature. Adhesion–time curves.

Thickness of aluminium 700 Å; thickness of chromium 460 Å.
High adhesion specimen, prepared using a shutter.

Acknowledgments

One of us (R.M.H.) wishes to acknowledge the assistance of the Cross Trust for the provision of a postgraduate research grant. We also wish to thank Mr. J. W. Sharpe for his assistance with the electron diffraction photographs.

References

American Society of Metals, 1948, *Metals Handbk.*
Bradley, A. J., and Lu, S. S., 1937, *J. Inst. Metals*, **60**, 319.
Gayler, M. L. V., 1937, *J. Inst. Metals*, **60**, 249.
Heavens, O. S., 1950, *J. Phys. Radium*, **11**, 335.
Heavens, O. S., and Collins, L. E., 1952, *J. Phys. Radium*, **13**, 658.
Hill, R. M., and Weaver, C., 1958, *Trans. Faraday Soc.*, **54**, 1464.
Holland, L., 1956, *Vacuum Deposition of Thin Films* (London: Chapman and Hall), p. 341.
Mott, N. F., 1947, *Trans. Faraday Soc.*, **43**, 429.
Smith, G. C., 1949, *Progr. Metal Phys.*, **1**, 163.
Townsley, M. G., 1945, *Rev. sci. Instrum.*, **16**, 143.
Weaver, C., and Benjamin, P., 1958 (in preparation).

Grain-Boundary Extrusion and Deformation during the Fatigue of Lead in a Partial Vacuum†

By K. U. Snowden
The Broken Hill Associated Smelters Pty. Ltd.
and The Baillieu Laboratory, University of Melbourne

[Received August 5, 1958]

Abstract

Observations have been made of slip, grain-boundary migration and grain-boundary extrusion during the fatigue of high purity lead *in vacuo* (approx. 5×10^{-3} mm Hg).

§ 1. Introduction

Recently, Forsyth (1957 a, b) has reported slip-band extrusion during fatigue of an aluminium + 7·5% zinc + 2·5% magnesium alloy, and Forsyth and Stubbington (1957) have observed grain-boundary extrusion during elevated temperature (250°C) fatigue of an aluminium + 4% copper alloy. In addition to slip-band extrusion, Cottrell and Hull (1957) have detected slip-band intrusion in copper fatigued at room and low temperatures. Extrusion and intrusion effects have been suggested, in theories by Cottrell and Hull (1957) and by Mott (1958), to be intimately associated with the start of fatigue cracks.

The following observations describe grain-boundary extrusions and deformation during the fatigue of lead *in vacuo* at room temperature.

§ 2. Experimental

Flat cantilever specimens with tapered sides were prepared from lead strip of 99·999% purity. The specimens were annealed at 100°C, chemically polished and tested at a small cyclic strain (estimated to be 0·05%) in reverse bending fatigue in a vacuum of the order of 5×10^{-3} mm Hg at a rate of 500 c.p.m.

§ 3. Results

Observations on the polished surface showed that the initial effect of cyclic straining caused slip and some boundaries to migrate. After about 10^4 to 10^5 cycles, the rate of migration slowed down and almost ceased. The initial and actual positions of a migrated boundary are shown in fig. 1‡ and are indicated by A and B respectively. After the rate of migration had diminished, specimens were lightly repolished to

† Communicated by Professor J. Neill Greenwood.
‡ All figures are Plates.

reveal more clearly the subsequent changes when given further cyclic straining. Figure 2 shows the same area as in fig. 1 after repolishing at $7{\cdot}5 \times 10^5$ cycles with $1{\cdot}25 \times 10^6$ additional cycles, i.e. a total of $2{\cdot}0 \times 10^6$ cycles. It is seen that the grain boundaries appear darkened and show little sign of further migration.

After repolishing and further cyclic straining slip tended to be denser in the central region of grains than in zones along either side of boundaries, fig. 2.

Examination of the grain-boundary zones at higher magnification showed that two types of extrusions with a metallic lustre had occurred along some boundaries. A whisker-like extrusion is shown in fig. 3 at 8×10^5 cycles and fig. 4 indicates the subsequent development of further whiskers, marked by arrows, in the same area after 2×10^6 cycles. The size of the whiskers varied and some were found up to 50μ long, with a diameter of approximately 1μ. Occasional ones had a fine structure —appearing twisted or kinked along their length.

A more continuous type of boundary extrusion is illustrated in fig. 5. This type often occurred together with the whisker-like extrusions.

In addition, some boundaries showed what appeared to be localized slip patterns which had geometrical forms, fig. 6. These were found on one or both sides of boundaries and their sides were frequently parallel to traces of slip systems in other parts of the same grain.

Further work is in progress.

Acknowledgments

This work forms part of a programme of research on the properties of lead, being carried on by the Broken Hill Associated Smelters in the Baillieu Laboratory, University of Melbourne. The work is being done under the direction of Professor J. Neill Greenwood, and the author gratefully acknowledges his help, together with the interest and encouragement of the Company's technical staff.

References

Cottrell, A. H., and Hull, D., 1957, *Proc. roy. Soc.* A, **242**, 211.

Forsyth, P. J. E., 1957 a, *Phil. Mag.*, **2**, 437; 1957 b, *Proc. roy. Soc.* A, **242**, 198.

Forsyth, P. J. E., and Stubbington, C. A., 1957, *J. Inst. Metals*, **85**, 339.

Mott, N. F., 1958, *Acta Met.*, **6**, 195.

The Stresses Produced in a Thin Elastic Plate by a Transverse Impulsive Force†

By C. J. Maiden
Engineering Laboratory, University of Oxford

[Received August 6, 1958]

Abstract

A description is given of tests in which a $\frac{1}{2}$ in. or $\frac{3}{8}$ in. diameter weighbar, seated vertically on a thin steel plate, was impacted by a finite mass. The force applied to the plate and the principal stresses produced at the bottom of the plate have been derived from the output of resistance strain gauges. The stress–time curves at the centre of the bottom surface of the plate, obtained at a number of impact velocities, have been compared with stress–time curves derived from theoretical considerations. The agreement between the theoretical and experimental results is found to be exceedingly good.

In such tests the plate showed no sign of having yielded even though the stresses reached were much greater than the static upper yield stress of the steel in the plate. This is attributed to the fact that strain rates of the order of 100 sec^{-1} were attained in the plate in these tests.

The manner in which the disturbance created in this type of test was propagated radially has been determined from the way the principal stresses, at the bottom of the plate, varied with radius and time. It is found that there was an annular area of compression which propagated radially in advance of a central area of tension.

Tests have also been made in order to determine the manner in which a disturbance caused by an impulsive point load propagates radially in a plate. A dural and a steel plate were impacted by a high-tensile steel sphere and the principal stresses produced at the bottom surface of the plates were once again determined from the output of strain gauges. It is found that the disturbances produced in these tests propagated in a similar manner to the way the disturbances propagated in the tests in which the force was applied over a finite area.

§ 1. Introduction

This investigation has been made in order to obtain information concerning the magnitude of the elastic stresses produced in a thin plate by a normal impulsive force. Tests are described in which a $\frac{1}{2}$ in. or $\frac{3}{8}$ in. diameter weighbar, seated vertically on a thin steel plate, was impacted by a finite mass, thus applying an impulsive load to the plate. The force–time curve applied to the plate in such a test has been obtained from the output of resistance strain gauges mounted on the bottom of the weighbar. Also the radial and circumferential stresses on the bottom of the plate have been derived from the output of strain gauges fixed to the plate surface.

† Communicated by the Author.

The stress–time curves at the centre of the bottom surface of the plate, one for each impact velocity concerned, have been compared with stress time curves derived from a theory presented by Sneddon (1945). Sneddon considered the bending of a plate caused by the application of a constant force uniformly distributed over a circle of radius a and acting for a finite time T. He derived expressions for the principal curvatures of the centre plane of the plate from which the stresses throughout the plate could be calculated. The results of the present tests have been compared with the theory by assuming that a particular force–time curve applied to the plate can be approximated by a number of superimposed square waves.

The manner in which the disturbance, created in such a test, has propagated radially has been investigated by plotting curves of radial and circumferential stresses against radius at a number of instants. Also, it was considered that a determination of the way in which a disturbance caused by an impulsive point load was propagated radially would be of interest. Hence tests have been made in which a dural and a steel plate have been impacted by a high-tensile steel sphere. The manner in which this type of disturbance has propagated in these plates has again been determined from the output of strain gauges fixed to the bottom surface of each plate.

§ 2. Apparatus

Figure 1 is a diagrammatic representation of the impact apparatus used in the present investigation. The cylindrical high-tensile steel mass, which impacted the weighbar, was free to slide in the head of a hammer arrangement. The length of the hammer, measured from the pivot to the centre of the steel insert, was 54 in. The dural hammer arm was of channel section and the head was machined from a cylindrical aluminium casting. The hammer was released from an electromagnet and arrested, after the steel insert had impacted the weighbar, by a rubber pad fixed to an aluminium angle which was supported at its ends so as to be above the plate. At the moment of impact the steel insert was horizontal and concentric with the weighbar.

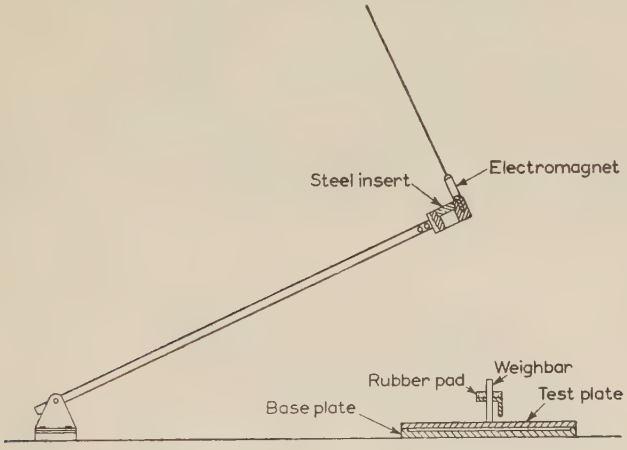
The plate used in such tests was 20·5 in. in diameter and 0·285 in. thick, and was bolted round the circumference to a levelled base plate. Two $3\frac{5}{8}$ in. long weighbars were used, one $\frac{1}{2}$ in. and the other $\frac{3}{8}$ in. in diameter; also the impact end of each weighbar was machined to a spherical shape of radius 10 in. For the steel used in the weighbars and plate, the elastic modulus was taken as 30×10^6 lb/in^2, the Poisson ratio as 0·30 and the density as 490 lb/ft^3.

Three wire resistance strain gauges were mounted on each weighbar. These gauges had a gauge length of $\frac{1}{4}$ in. and were centred $\frac{1}{4}$ in. from the plate end of each weighbar. Also strain gauges were fixed to the bottom surface of the plate at the centre and at radii of 1, 2, 3, 4, 5 and 6 in. from the centre respectively. At each of these radii there were two gauges,

one in a radial direction and the other in a circumferential direction. Only one gauge was required at the centre of the plate, for at any given instant the radial and circumferential strains were equal at this point.

The electronic apparatus used to record the outputs from these gauges was almost the same as that used in an earlier investigation (Campbell and Maiden 1957). In these previous tests the cathode ray oscilloscope spot was triggered and brightened by the magnetostrictive output of a coil surrounding a weighbar. In the present tests the spot was triggered and brightened by the gauge output itself.

Fig. 1



Mechanical impact apparatus.

§ 3. Theoretical Solution for a Plate Impacted by a Constant Force

Let ρ, E and σ refer to the density, elastic modulus and Poisson ratio respectively of the material in a plate of thickness $2h$. Also let $D=2Eh^3/3(1-\sigma^2)$ denote the flexural rigidity of the plate.

Boussinesq (1885) has shown that if a normal force is applied at a point on a large elastic plate, then at any instant the velocity v of that point is proportional to the force f acting on the plate. He found that the relationship between these two quantities was given by

$$f=16\rho hbv \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where $b^2=D/2\rho h$.

Sneddon has used this relationship in deriving, from the differential equation governing the transverse displacement of a thin plate, expressions for the principal curvature at the centre of a plate when loaded by a constant force F applied over a circle of radius a and acting for a time T. The expressions for the curvature κ at a time t are given by

$$\kappa=-[\log_e(a^2/4bt)+\gamma-1]F/16\rho\pi hb^2 \quad \text{for } 0\leqslant t\leqslant T \quad . \quad . \quad . \quad (2)$$

and

$$\kappa=-[\log_e(1-T/t)]F/16\rho\pi hb^2 \quad \text{for } t\geqslant T \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where γ is Euler's constant. It is to be noted that this formula is accurate only if $a^2/4bt$ is small.

Also, if M is defined as the bending moment per unit length at the centre of the plate then

$$M = D\kappa(1+\sigma) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

and the tensile stress p at the bottom surface of the plate can be determined from the formula

$$p = 1{\cdot}5M/h^2. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Hence from eqns. (2), (3), (4) and (5) the stress p, at any time t, is given by

$$p = -[\log_e(a^2/4bt) + \gamma - 1]F(1+\sigma)/5{\cdot}33\pi h^2 \quad \text{for } 0 \leqslant t \leqslant T \quad . \quad . \quad (6)$$

and

$$p = -[\log_e(1 - T/t)]F(1+\sigma)/5{\cdot}33\pi h^2 \quad \text{for } t \geqslant T. \quad . \quad . \quad . \quad . \quad (7)$$

This solution only applies when a plate is subjected to a force of constant magnitude, a condition which would not often be met in practice. In the next section it will be shown how the theory can be used when a force–time curve of general shape is applied to a plate.

§ 4. Experimental Results

4.1. *The Magnitude of the Tensile Stress Produced Immediately Below the Load*

The $\frac{1}{2}$ in. diameter weighbar, seated vertically on the plate, was impacted at four different hammer head velocities. At each impact velocity four strain–time records were taken from the weighbar gauges and four from the gauge at the centre of the plate; the weighbar records were found to be repeatable to within $\pm 3\%$ and the records from the plate gauge to within $\pm 2\%$.

A typical plate strain–time oscillogram, with calibration lines superimposed during enlarging, is presented in fig. 2 (*a*), Pl. 101. All the traces were interrupted every two microseconds in order to give an accurate time base.

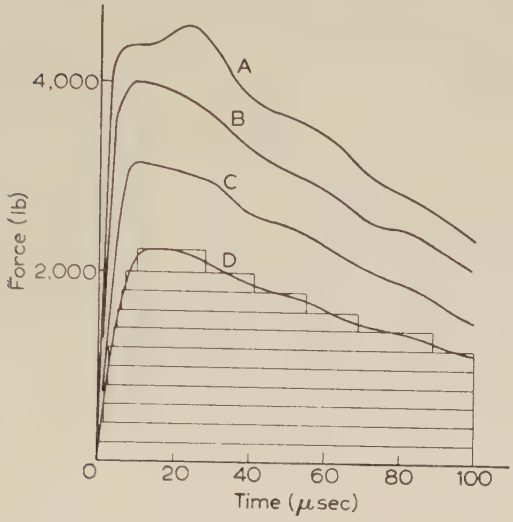
The force–time curve applied to the plate, at a particular impact velocity, was determined by multiplying the relevant average weighbar strain–time record by $\pi a^2 E$, where E is the elastic modulus for the weighbar steel. The derived force-time curves, one for each of the four impact velocities concerned, are shown in fig. 3 (*a*).

The stress–time curve at the centre of the bottom surface of the plate, for each impact velocity, was determined from the relevant average strain–time record using the formula relating the stress p and strain ϵ at the centre of a plate. This formula is given by

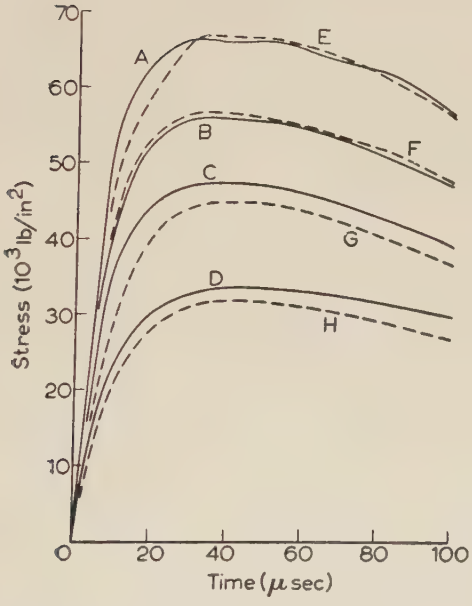
$$p = E\epsilon/(1-\sigma). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

The stress–time curves obtained in this manner are shown in fig. 3 (*b*).

Fig. 3



(*a*)



(*b*)

Force–time and stress–time curves obtained at various heights of drop using a $\frac{1}{2}$ in. diameter weighbar.

(*a*) Force–time curves. Curves A, B, C and D correspond to heights of drop of 37·4, 28·0, 22·8 and 10·2 in. respectively.

(*b*) Stress–time curves. Curves A, B, C and D refer to experimental results and curves E, F, G and H to theoretical results at height of drop of 37·4, 28·0, 22·8 and 10·2 in. respectively.

The force–time curves applied to the plate in these tests were by no means square waves. Hence, in order to compare Sneddon's theory with the curves shown in fig. 3 (*b*), each force–time curve has been considered to be equivalent to a number of superimposed square waves. The magnitude of all these square waves, excluding the top increment, has been taken as 200 lb. Curve D in fig. 3 (*a*) shows such an approximation. For each force–time curve eqns. (6) and (7) have been used to determine the effect of each increment of load on the stress at the centre of the plate. Thus, by adding all the increments of stress at corresponding times, the theoretical stress–time curve has been found at each impact velocity. The four theoretically derived stress–time curves are presented in fig. 3 (*b*) (broken curves), and it is seen that they compare very favourably with the experimental stress–time curves. It is to be noted that eqns. (6) and (7) only apply to a plate of infinite radius, whereas the plate used in the present investigation had a radius of 10·25 in. In the next section it will be shown that, for the time interval over which the theory has been compared with the experimental results, the centre of the plate acted as though the plate was infinitely large.

It is of interest to note that the thickness of the plate was chosen so that the plate should theoretically act as an extension of the weighbar. This has been arranged in the following manner. Provided there is no wave reflection at the weighbar-plate interface then the force f applied to the plate at any instant is given by $\rho c v \pi a^2$, where c is the velocity of sound in the weighbar. This force is also given by $16\rho h b v$, hence equating these two values for f we obtain

$$16hb = c\pi a^2$$

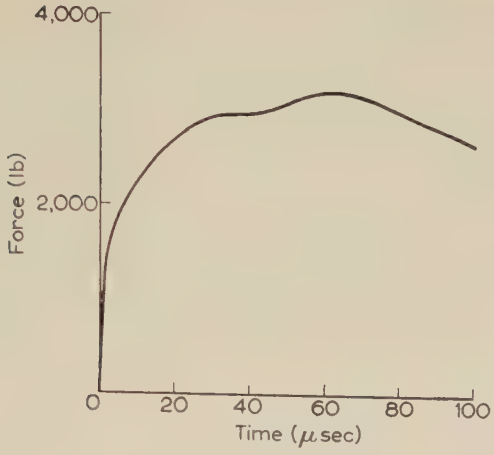
i.e.
$$h = \pi^{1/2} 3^{1/4} a (1-\sigma^2)^{1/4}/4 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

which for a $\frac{1}{2}$ in. diameter weighbar yields $h = 0{\cdot}1425$ in. That the actual reflected waves were small is seen from the face that the oscillations on the force–time curves are slight; and also from the fact that the force–time curves die off approximately exponentially, this being a theoretical condition for a finite mass striking a semi-infinite rod.

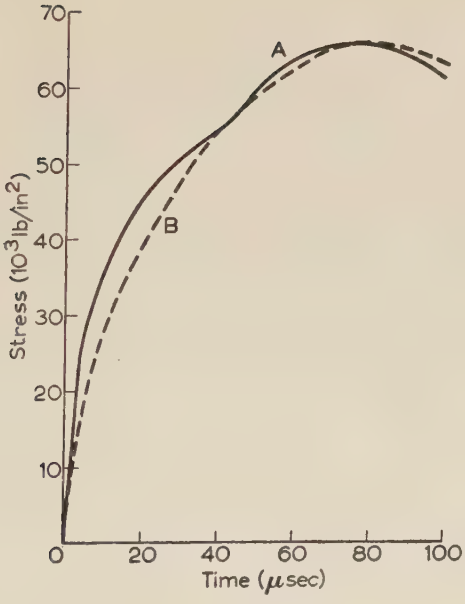
Tests were also made using a $\frac{3}{8}$ in. diameter weighbar and the results obtained have been compared with the theory. Only one impact velocity was considered, this corresponding to a 28 in. height of drop. The average force–time curve applied to the plate is shown in fig. 4 (*a*) and the average stress–time curve, at the centre of the bottom surface of the plate, is presented in fig. 4 (*b*). Once again the force–time curve was considered to be composed of a number of square waves in order to derive, from eqns. (6) and (7), the theoretical stress–time curve at the centre of the plate. This theoretical curve is shown in fig. 4 (*b*) and the agreement with the experimental curve is again seen to be excellent.

In all the above tests the resistance, measured to five figures, of the strain gauge at the centre of the plate, was found to be the same after test as before test. It was thus assumed that in all tests the plate remained elastic.

Fig. 4



(*a*)



(*b*)

The force–time and stress–time curves obtained using a 28 in. height of drop and a $\frac{3}{8}$ in. diameter weighbar.

(*a*) Force–time curve.

(*b*) Stress–time curves. Curves A and B refer to the experimental and theoretical results respectively.

4.2. *Propagation of a Disturbance Radially*

4.2.1. *Impact over a finite area*

It was described in § 2 how gauges had been placed in both radial and circumferential directions at radii of 1, 2, 3, 4, 5 and 6 in. from the centre of the plate. In the present tests the $\frac{1}{2}$ in. diameter weighbar was impacted at a velocity corresponding to a 28 in. height of drop and oscillograms were obtained giving the output from the central gauge plus that from each of the above twelve gauges in turn. The known output of the central gauge was subtracted from these dual outputs in order to obtain the strain–time curves, both radial and circumferential, at each of the six radii considered. In this way the time base for all the strain–time curves was the same. Figure 2 (*b*), Pl. 101, shows a typical strain–time oscillogram obtained in the above manner.

At each radius the radial and circumferential stress–time curves have been found from the radial and circumferential strain–time curves by using the formulae relating the principal stresses (p_r and p_c) to the principal strains (ϵ_r and ϵ_c) at a point on a plate. These formulae are given by

$$p_r = (\epsilon_r + \sigma\epsilon_c)E/(1-\sigma^2) \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

and

$$p_c = (\epsilon_c + \sigma\epsilon_r)E/(1-\sigma^2) \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

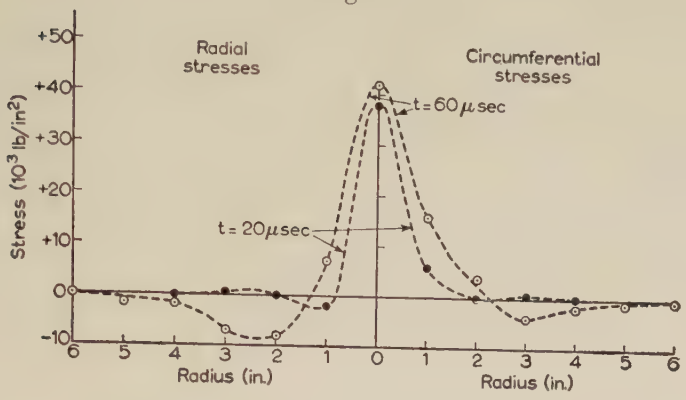
where the subscripts r and c refer to the radial and circumferential directions respectively. From the derived curves of radial stress–time and circumferential stress–time, the radial and circumferential stresses have been plotted against radius at times of 20 and 60 microseconds respectively. These curves are shown in fig. 5 (*a*).

It is to be observed that up to 60 microseconds the disturbance had not significantly effected points at any radius greater than about 6 in. Hence it is considered that, for the time interval over which the theory has been compared with the experimental results presented in the previous section, the centre of the plate behaved as though the plate was infinitely large.

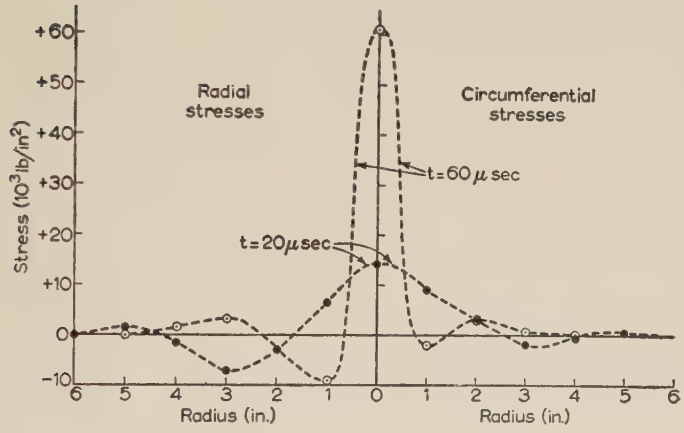
4.2.2. *Impact at a point*

It was decided to investigate the way in which a disturbance caused by an impulsive point load was propagated radially in a plate. A reasonable approach to this type of loading arises when a plate is impacted by an elastic sphere; hence tests have been made in which two plates, one dural and the other steel, were impacted by a 1 in. diameter high-tensile steel sphere at a velocity of 243 in/sec. The dural and steel plates were 0·133 in. and 0·330 in. in thickness respectively. In order to prevent localized yielding at the point of impact in such tests two methods, one for each plate, were used. For the dural plate the sphere impacted a small shim of dural, 1 in. in diameter and $\frac{1}{32}$ in. thick, placed on the centre of the plate; and for the steel plate impact occurred at the centre of a

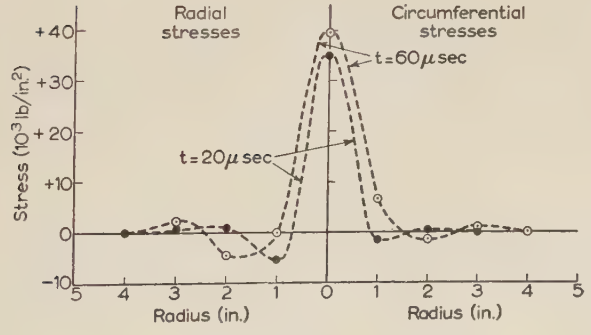
Fig. 5



(*a*)



(*b*)



(*c*)

Curves of radial and circumferential stresses against radius at times of 20 and 60 μsecs. Tension $+ve$; Compression $-ve$.

(*a*) Tests on steel plate (0·285 in.) using $\frac{1}{2}$ in. diameter weighbar. Height of drop: 28 in.

(*b*) Tests on steel plate (0·330 in.) using 1 in. diameter steel sphere. Impact velocity: 243 in./sec.

(*c*) Tests on dural plate using 1 in. diameter steel sphere. Impact velocity: 243 in./sec.

high-tensile steel insert, 1 in. in diameter and $\frac{1}{8}$ in. thick, which had been pressed into a similar sized cavity machined out of the centre of the upper surface of the plate. For each plate the radial and circumferential stress–time curves at the centre and at radii of 1, 2, 3, 4, 5 and 6 in. were derived from the output of strain gauges in a similar manner to that described in the previous section. In using eqns. (10) and (11) to derive the curves for the dural plate, E was taken as 10×10^6 lb/in^2 and σ as 0·34. For both plates the radial and circumferential stresses have again been plotted against radius at times of 20 and 60 microseconds respectively; these curves are shown in figs. 5 (*b*) and 5 (*c*). A typical strain–time oscillogram used in the derivation of these curves is presented in fig. 2 (*c*), Pl. 101.

§ 5. Discussion and Conclusions

The agreement between Sneddon's theory and the results of the tests using the hammer impact apparatus is exceedingly good. With reference to figs. 3 (*b*) and 4 (*b*) the discrepancy, over the first few microseconds, between each theoretical and experimental curve can be partially explained by the fact that $a^2/4bt$ must be small for eqns. (6) and (7) to be accurate. For the $\frac{3}{8}$ in. and $\frac{1}{2}$ in. diameter weighbars the magnitudes of $a^2/4bt$ at a time of 10 microseconds are 0·0512 and 0·0912 respectively; hence it is considered that only for times less than this would eqns. (6) and (7) be in significant error. Three other factors must also give rise to differences between the theoretical and experimental results. The first factor is that the applied force–time curves had to be approximated to in order to calculate the theoretical curves; the second is that the effects of ' rotary inertia ' and shear forces are neglected in the theory; and the third factor is that the effects of local elastic deformation at the weighbar-plate interface have not been taken into account in deriving the theoretical stress–time curves. However, the difference between the theoretical and the experimental results is considered to be small enough to justify the use of the theory in the design of plates to sustain transverse impact loads. This could be done provided the impact velocity, the area over which the load is applied and the duration of the impact are known. The maximum tensile stress in the plate could then be calculated from eqn. (6) if it is assumed that the force–time curve applied to the plate is a perfect square wave.

Also it is to be noted that in the tests made, using both the $\frac{1}{2}$ in. and $\frac{3}{8}$ in. diameter weighbars, the plate showed no sign of having yielded. The maximum tensile stress reached during these tests was 66 500 lb/in^2 which is well above the static upper yield stress (35 000 lb/in^2) found for the steel used in the plate. This increase in strength is to be expected due to the fact that during these tests strain-rates of the order of 100 sec^{-1} were attained in the plate. Impact compression tests (see Campbell and Duby 1957) were made on specimens of the same steel as used in the plate in order to give an indication of the increase in strength of the steel at

such high strain rates. It was found that for a pre-yield strain-rate of the order of 100 sec^{-1} the upper yield stress was raised to approximately 85 000 lb/in^2.

It is of interest to calculate the tensile stresses that would be reached at the centre of the bottom surface of the plate if it were loaded statically by each of the maximum forces given in fig. 3 (*a*), and to compare these stresses with the actual maximum tensile stresses produced in the plate. This has been done as follows. Woinowsky-Krieger (1933) has shown that for a plate of radius R, clamped round the edge, the tensile stress p immediately below the load P is given by

$$p = P(1+\sigma)[0{\cdot}485 \log_e (R/2h) + 0{\cdot}52]/4h^2 \quad . \quad . \quad . \quad . \quad (12)$$

where p is independent of the radius a of the area of loading provided $a \ll R$. Now fig. 3 (*a*) shows that, for the four impact velocities considered, the maximum loads applied to the plate were 4540, 4000, 3160 and 2240 lb. Substituting these values for P in eqn. (12) the corresponding tensile stresses are found to be 165 000, 145 000, 114 000 and 81 000 lb/in^2 respectively. A comparison of these values with the actual maximum tensile stresses produced at the four impact velocities, i.e. 66 500, 56 000, 47 500 and 33 500 lb/in^2, show that eqn. (12) is quite inadequate for dealing with impulsive loads.

Figures 5 (*a*)–(*c*) show that, for the steel plate (0·285 in.) impacted over a finite area and for the steel (0·330 in.) and dural plates impacted at a point, the radial and circumferential tensile stresses produced at the bottom surface of each plate were still very localized even after 60 microseconds. Also it is to be observed that in all three tests there was an area of compression which propagated radially in advance of the central area of tension. These results are similar to those that would be predicted from the curves of plate deflection against time derived by Sneddon for the particular case of a square loading wave applied to a plate.

Acknowledgment

The author is grateful to Dr. J. D. Campbell for the advice he has so generously given throughout the course of this investigation.

References

Boussinesq, J., 1885, *Application des potentials . . .*, 464.
Campbell, J. D., and Duby, J., 1957, *Conference on Properties of Materials at High Rates of Strain* (Inst. Mech. Engrs).
Campbell, J. D., and Maiden, C. J., 1957, *J. mech. phys. Solids*, **6**, 53.
Sneddon, I. A., 1945, *Proc. Camb. phil. Soc.*, **41**, 27.
Woinowsky-Krieger, S., 1933, *Ingen-Arch.*, **4**, 305.

The Interpretation of Soft X-Ray Emission Spectra

By J. A. Catterall and J. Trotter
Communication from the National Physical Laboratory

[Received August 15, 1958]

Abstract

An examination of the K satellite emission bands observed in lithium and beryllium has indicated that the presence of the hole which occurs in the core-level during the soft x-ray emission process is unlikely appreciably to affect the intensity distribution of the spectra.

§ 1. Introduction

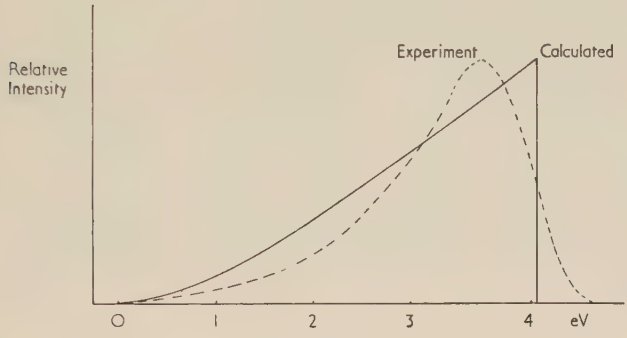
The method of soft x-ray spectroscopy has been used for a number of years as a means of investigating the electron energy distribution of the valence bands of solids. Reviews of the subject have been given by Skinner (1946) and Tomboulian (1957). Much of the experimental work has been carried out on metals and alloys and it is with the soft x-ray bands of metals that this paper is concerned.

Until recently it was accepted that the soft x-ray emission spectra gave a general indication of the distribution with energy of the various types of electron states in the valence band. That is to say, K spectra gave a measure of the density of states of p symmetry, and L spectra the density of states of $s+d$ symmetry. However, recent calculations of the band structure of lithium (Schiff 1954, Parmenter 1952, Brown and Krumhansl 1958, Glasser and Callaway 1958) which show a predominance of p states at the Fermi surface, are at variance with the observed K spectrum which shows a maximum before the high energy end of the band is reached (fig. 1). Jones and Schiff (1954) have pointed out that since x-ray emission takes place at an atom which has already lost one electron the states in the neighbourhood may be strongly perturbed. Skinner *et al.* (1954) do not agree with this view since Schiff's (1954) calculated bandwidth is considerably different from the experimental value. In the present paper an attempt is made to assess the magnitude of the effect of the loss of an electron on the band shapes and widths by considering these properties in the satellite spectra.

Soft x-ray emission bands are produced when an electron is ejected from an inner shell of an atom and the resulting hole is filled by a transition from the band of valence electrons. The intensity is given by the expression $I(E)=\nu^3 \, . \, F(E) \, . \, N(E)$ where ν is the frequency, $F(E)$ the

transition probability, and $N(E)$ the density of states. However, in addition satellite bands are frequently observed at higher energies, arising from double ionization of the inner shells. In this case the radiation corresponds to a final state in which there is a hole in the valence band and a hole also remaining in an inner shell. If the effect of a positive hole is large one would expect a considerable difference between the shapes and widths of main bands and satellite bands, and a comparison between the two enables an estimate to be made of the effect of the hole on the product $F(E) \cdot N(E)$. In comparing these bands it is necessary to allow for the broadening due to the shorter lifetime of the doubly-ionized state.

Fig. 1



It is the K spectra which are of primary importance for comparison of the bands. The unexcited K level has a spectroscopic designation $^1S_0$. Removal of one electron converts it to the state $^2S_{1/2}$, whilst removal of a second electron results once more in the state $^1S_0$. Owing to the selection rules the K spectra represent the p fraction of the valence (or conduction) band. The main emission corresponds to transitions down to the $^2S_{1/2}$ state, and the satellite emission to transitions to the $^1S_0$ state. Since in both types of transition the lower level is a singlet a comparison between the two types of emission may be made directly. L satellite spectra are much more difficult to interpret. Removal of one electron from the 2s state gives the term $^2S_{1/2}$, whilst removal of one from the 2p state gives the terms $^2P_{1/2}$ and $^2P_{3/2}$. However, since emission bands corresponding to transitions down to the $^2S_{1/2}$ level are not observed, due to an Auger process which fills the 2s hole, we may disregard ionization in the 2s shell and consider only the removal of 2p electrons. Removal of a second 2p electron gives the terms $^3P_2$, $^3P_1$, $^3P_0$, $^1D_2$ and $^1S_0$. The L satellite spectra consist therefore of transitions between the p fraction of the valence band and the levels $^1D_2$ and $^1S_0$, and the $s+d$ fraction and the levels $^3P_2$, $^3P_1$ and $^3P_0$. Calculations by Kennard and Ramberg (1934) of the energy separation between the above five levels show that the two types of satellite band must overlap. Consequently interpretation of L

satellites is quantitatively very difficult, although an attempt at a qualitative interpretation of the bands obtained by Skinner has been made here, in order to show that the general features are those that would be expected.

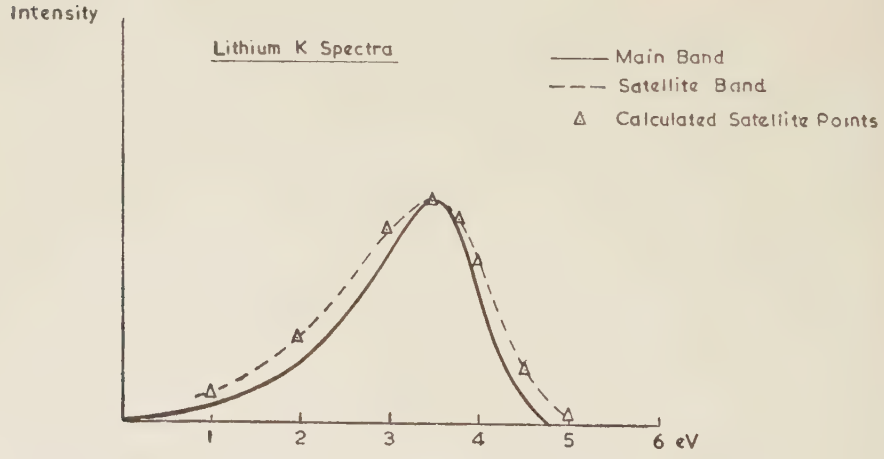
§ 2. Experimental Technique

The spectrograph in use at the National Physical Laboratory and the technique used in its operation have been described elsewhere (Catterall *et al.* 1958).

§ 3. K Satellites of Li and Be

The K emission main band of lithium has been measured by Skinner (1946), by Bedo and Tomboulian (1958) and by Catterall and Trotter (1958). The general features of the band are in agreement. In addition Tomboulian has measured the satellite band of lithium, with which we also agree. The intensity of the satellite is much lower than that of the main band, and occurs at 88 ev, whereas the main band is at 55 ev. In fig. 2 are shown the main and satellite bands superimposed to match at the peak intensities. It will be seen that the two bands are very similar except for an overall broadening of the satellite. We have also measured the K satellite band of beryllium which occurs at 134 ev, and it is shown in fig. 3.

Fig. 2

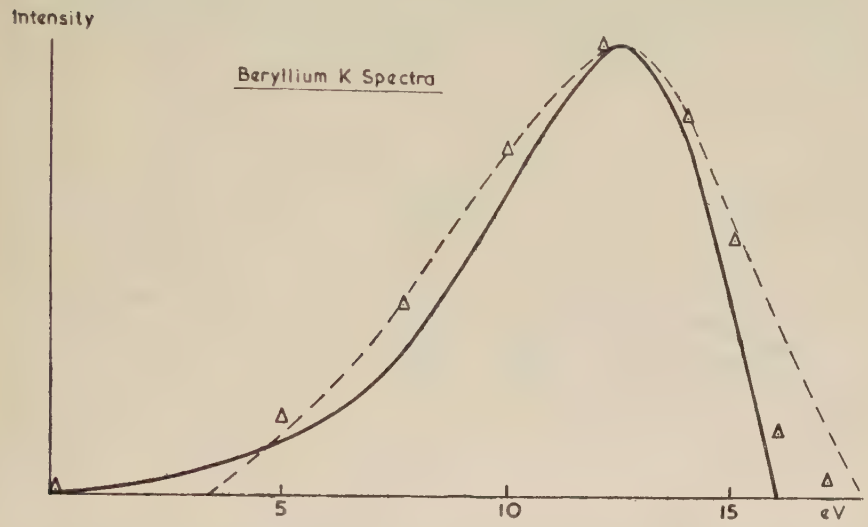


The emission intensity distribution of both main and satellite bands depends not only upon the electron energy distribution in the valence band of the solid but also upon the ' natural width ' of the core level. Theoretically the line shape of the core level should be Lorentzian,

$$I = \frac{I_0}{a^2 + (E - E')^2}$$

with a width at half-maximum intensity of $2a$. Unfortunately the half widths of levels for elements with atomic numbers as low as lithium and beryllium are not known. However, from a study of the width of the K absorption limit in lithium Bedo and Tomboulian (1958) conclude that the K level width may be about 0·1 ev. Using the empirical rule that the

Fig. 3



half-width of a line increases with the square of the excitation energy, the half-width of the doubly-ionized state should be about 0·3 ev in the case of lithium and 0·7 ev in the case of beryllium. It is possible to calculate the expected shape of the satellite band from a knowledge of the main emission band and the natural width of the doubly-ionized level. The calculated shape will only be approximate since the main band will itself be broadened by the natural width of the singly-ionized level. However, in figs. 2 and 3 are the points calculated by a ' folding ' operation, which involves determining the integral

$$I_2(E) = \int_{-\infty}^{+\infty} I_1(E') \cdot \frac{1}{a^2 + (E - E')^2} dE'$$

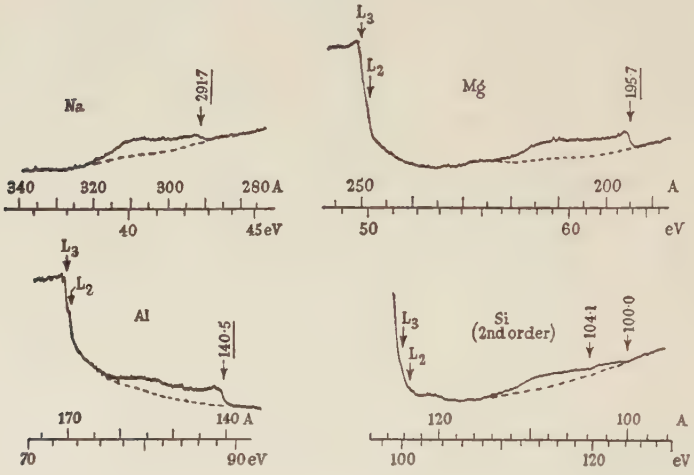
where $I_2(E)$ is the intensity of the satellite band, and $I_1(E')$ the intensity of the main band. The agreement between the observed and calculated satellite points is quite good. The fit at the low intensity parts of the beryllium satellite is only fair, probably because the beryllium satellite is more difficult to measure since its intensity is very low.

§ 4. L Satellites of Na, Mg, Al and Si

The bands discussed in this section are those obtained by Skinner (1946) and shown in fig. 4. No detailed analysis of the shape of the L satellite bands is possible because of the splitting of the L level with the removal

of the second electron. The separation in electron volts between the level $^3P_2$ and the others, calculated by Kennard and Ramberg for isolated atoms, is shown for the four elements in the table. The $^1S_0$ level has the highest energy. The relative intensity of the emission from the valence band (V) to these levels should be given by their quantum weight, and should be in the ratio $^3P_{210} : {}^1D_2 : {}^1S_0 = 9 : 5 : 1$. The intensity of the satellite bands is usually only 5% or so of that of the main bands, and in view of the greater experimental difficulties in the measurements we may neglect the level $^1S_0$ and consider the satellites to be made up from the

Fig. 4



	Na	Mg	Al	Si
$^3P_2$	0	0	0	0
$^3P_1$	0	0·14	0·27	0·34
$^3P_0$	0	0·18	0·54	0·68
$^1D_2$	3·1	3·87	4·5	5·1
$^1S_0$	8·1	9·64	11·2	12·6

transitions $V \rightarrow {}^1D_2$ and $V \rightarrow {}^3P_{210}$. The transition $V \rightarrow {}^1D_2$ has the highest energy and will be responsible for the high energy edge of the band, whilst the transition $V \rightarrow {}^3P_{210}$ will overlap at lower energies. The selection rules indicate that the former transition will arise from the p fraction of the valence band, and the latter from the $s+d$ fraction. The intensity distribution given by the p fraction is the same as in the K spectra of main emission bands, and the $s+d$ fraction the same as in the L spectra. Accordingly, a qualitative analysis of the shape of the satellite bands

should show that the high energy end has the same features as the main K spectra, with a somewhat broader edge, and with a more intense band resembling the main L spectra overlapping it at an energy lower by an amount equal to the $^3P_{210}$-$^1D_2$ separation. The two overlapping bands may not have their intensities in the precise ratio 9:5 since the rule relating the relative intensities requires a common initial level before the transition, whereas in the present case the initial levels in the valence band are different. The most that can be expected is that the satellite bands should show a rise in intensity at the lower energies compared with the low energy end of the K spectra.

4.1. *Sodium*

There are no published data concerning the shape of the sodium spectrum, but the theoretical calculations of the valence band show that there is a predominance of *s* states at the Fermi surface (Schiff 1954) so that the K spectrum should decrease before the high energy edge is reached. Consequently no sharp features such as a peak will be expected at the high energy edge of the satellite band (although there should be some broadening of the edge) and none are observed. The satellite shows the expected broad hump at about 2·5–3·0 ev from the edge. The $^3P_2$–$^1D_2$ separation in sodium is 3·1 ev.

4.2. *Magnesium*

The K spectrum reported by Farineau (1937) has a sharp peak at the high energy edge (not unlike that in the L spectrum). A peak is also present at the satellite band edge, and a broad hump occurs at an energy lower by about 4·5 ev. The $^3P_2$–$^1D_2$ separation in magnesium is 3·87 ev.

4.3. *Aluminium*

The K spectrum obtained by Farineau (1936) is illustrated by Skinner (1946), and like magnesium shows a sharp peak at the high energy edge. The satellite also shows a peak, with a very broad hump at about 7 ev lower. The $^3P_2$–$^1D_2$ separation is 4·5 ev.

4.4. *Silicon*

The K spectrum of silicon obtained by Farineau (1937) has no sharp emission edge but decreases from a maximum to zero over a range of 4 ev. The satellite band edge shows the same general feature. Skinner reports the satellite band as extending to 100·0 Å, and there is a broad hump at a lower energy some 10 ev from this point. The band V→$^3P_{210}$ seems to begin at the point 104·1 Å where Skinner records a change of slope in the emission curve, and this is 4·9 ev from the high energy end. The $^3P_2$–$^1D_2$ separation in silicon is 5·1 ev.

The interpretation of the silicon satellite is further complicated by the fact that in insulators and semiconductors, unlike metals, an excess

positive charge on an ion will definitely produce a bound state lying below the valence band (Koster and Slater 1954). In silicon this will be a 3s state, and transitions from this level to the lower levels will give a broad line on the low frequency side of the spectrum. The energy difference between this line and the high energy end of the satellite spectrum will be approximately the same as the $(3s)^2 \rightarrow (3s)(3p)$ transition in the element of next highest atomic number P^{4+}. This is 8·4 ev for parallel spin of the 3p screening electron. The transition accordingly may account for the broad hump appearing some 10 ev from the high energy end of the satellite.

§ 5. Discussion and Conclusions

Examination of the shape of the satellite bands of lithium and beryllium shows them to be identical with those of the main band, except for a broadening which can be explained by the natural width of the core level. In addition, a simple qualitative interpretation of the shapes of the L satellite bands seems possible by considering the splitting of the L level and the shapes of the K and L main band spectra.

We may consider separately the possible effect of the extra positive hole on (*a*) the electron energy distribution $N(E)$ and (*b*) on the transition probability $F(E)$ respectively.

(*a*) The excess positive charge due to the hole remaining in the inner level will be screened either by a level subtracted from the valence band in the manner of Friedel (1952) (although Slater (1955) considers this unlikely with only one excess positive charge) or by the electrons in the valence band itself. In the first case a broad line may be visible, if the transition is allowed, on the low frequency side of the band, and in the second case the band itself may be somewhat altered in shape.

(*b*) The transition probability $F(E)$ for the main emission bands depends upon the wave functions of the initial and final states. The extension of the wave functions in space will be changed by the increased effective nuclear charge due to the presence of the positive hole, and consequently we may expect a difference between $F(E)$ in satellites and main bands.

Since the extra positive hole, however, has no apparent effect on either $F(E)$ or $N(E)$ we may reasonably conclude that the hole in the core level, which appears during the main band emission process, does not appreciably alter either of these terms. We may also conclude, that the suggestion of Jones and Schiff, that the presence of the positive hole could account for the large discrepancy between the observed and calculated band forms of lithium, is extremely improbable.

Acknowledgment

The work described above has been carried out as part of the General Research Programme of the National Physical Laboratory and this paper is published by permission of the Director of the Laboratory.

References

Bedo, D. E., and Tomboulian, D. H., 1958, *Phys. Rev.*, **109,** 35; also *ONR Technical Report*, February 1957.

Brown, E., and Krumhansl, J. A., 1958, *Phys. Rev.*, **109,** 30.

Catterall, J. A., and Trotter, J., 1958 (to be published).

Catterall, J. A., Wilson, L. F., and Trotter, J., 1958, *J. sci. Instrum.* (in the press).

Farineau, J., 1936, *C.R. Acad. Sci., Paris*, **203,** 540; 1937, *Ibid.*, **204,** 1108.

Friedel, J., 1952, *Phil. Mag.*, **43,** 153.

Glasser, M. L., and Callaway, J., 1958, *Phys. Rev.*, **109,** 1541.

Jones, H., and Schiff, B., 1954, *Proc. phys. Soc. Lond.* A, **67,** 217.

Kennard, E. H., and Ramberg, E., 1934, *Phys. Rev.*, **46,** 1040.

Koster, G. F., and Slater, J. C., 1954, *Phys. Rev.*, **96,** 1208.

Parmenter, R. H., 1952, *Phys. Rev.*, **86,** 552.

Schiff, B., 1954, *Proc. phys. Soc. Lond.* A, **67,** 2.

Slater, J. C., 1955, *Phys. Rev.*, **98,** 1039.

Skinner, H. W. B., 1946, *Phil. Trans.* A, **239,** 95.

Skinner, H. W. B., Bullen, T. G., and Johnston, J. E., 1954, *Phil. Mag.*, **45,** 1070.

Tomboulian, D. H., 1957. *Handbuch der Physik* (Berlin: Springer), vol. 30, p. 246.

The Dynamic Compression of Perspex†

By J. S. Buchanan, H. J. James and G. W. Teague
Ministry of Supply,
Armament Research and Development Establishment, Fort Halstead

[Received July 25, 1958; and in revised form August 27, 1958]

Abstract

High explosives have been employed to generate intense shock waves in Perspex and from the measurements made of the shock and free surface velocities the ' Hugoniot ' equation of state has been determined within the pressure range from 3 to 13×10^{10} dynes cm^{-2}. The results fit an equation of an inverted Bridgman type. Graphs are given of the variation of pressure, shock and particle velocity with Perspex thickness.

§ 1. Introduction

The detonation of high explosives in contact with a solid results in the transmission of a steep fronted pressure pulse into the material. The pressures obtained by this means are greater than the highest pressures reached in static experiments such as carried out by Bridgman (1945 a, b, c, 1949). The propagation of these high intensity shock waves in metals is a subject which has received considerable theoretical attention in the last few years (Duval and Zwolinski 1955, Walsh and Christian 1955, Goranson *et al.* 1955) and it has been shown that an equation of state can be derived from simultaneous measurements of the velocity of the shock wave and the particle velocity (mass velocity) in the wave. The pressure volume relations (often called Hugoniots) obtained from experiments of this nature are in a pressure region so far unattainable by other methods. These dynamic experiments also have the advantage that they avoid the use of cumbersome apparatus, with its inherent difficulties of unknown relative distortions which lead to large uncertainties in measurements. The results from the two types of experiments are not immediately comparable since in one case the pressure is applied very slowly and under isothermal conditions, while in the other the pressure is applied extremely rapidly ($< 10^{-6}$ sec) with an associated rise in temperature. Nevertheless the difference between the two cases is not large and it is possible to convert from one to the other by the use of thermodynamic relations (Goranson *et al.* 1955).

Much experimental work has been carried out along these lines since the pioneer experiments of Pack *et al.* (1948). Their original technique has been developed (Minshall 1955, Bancroft *et al.* 1956), optical techniques have been used (Walsh and Christian 1955, Walsh and Rice 1957) and as a result a large amount of data is available on the

† Communicated by the Authors.

compressibility of metals (Walsh and Christian 1955, Minshall 1955, Bancroft *et al.* 1956, Walsh *et al.* 1957). Some work has been carried out on liquids (Walsh and Rice 1957, Rice and Walsh 1957) and air (Deal 1957), but very little has been reported about the transmission of these shock waves through non-metallic solids (Lawton and Skidmore 1956). Due to the increasing utilization of plastics it was considered desirable to obtain quantitative information about their behaviour. Perspex was chosen since it is widely used, readily available and easily machined, which facilitated the preparation of the test assemblies.

§ 2. Theory

The equations based on the conservation of mass and momentum across a plane shock wave were derived by Rankine and Hugoniot and are respectively,

$$\frac{\rho}{\rho_0} = \frac{D}{D-u} \qquad \ldots\ldots\ldots \quad (1)$$

and

$$P - P_0 = \rho_0 u D \qquad \ldots\ldots\ldots \quad (2)$$

where P, ρ and u are the pressure, density and particle velocity in the shock wave which is travelling with velocity D relative to the material at rest, with initial density ρ_0 at pressure P_0. P_0 is usually taken as zero when $P \gg P_0$.

The shock velocity, D, can be measured directly but the particle velocity, u, in a shock wave must be deduced from the measurement of the free surface velocity which results from the normal reflection of the shock wave at the free surface. There must be zero pressure difference across the free surface boundary, and this condition requires that a wave equal in amplitude but opposite in sign must be reflected from the surface. Under these conditions the surface will acquire a velocity

$$v = u + \sigma \qquad \ldots\ldots\ldots \quad (3)$$

where u is the particle velocity in the stress wave, and σ is the excess of the free surface velocity over the particle velocity u. If it is assumed that the unloading takes place isentropically then σ is the Riemann function and is given by

$$\begin{aligned} \sigma &= \int_{\rho}^{\rho_0} \frac{c}{\rho}\, d\rho \\ &= \int_{0}^{P} \left(-\frac{\partial v}{\partial p} \right)_s^{1/2} dp \qquad \ldots\ldots\ldots \quad (4) \end{aligned}$$

where subscript s denotes constant entropy, and c is the velocity of sound.

The approximate relation

$$u = \sigma \qquad \ldots\ldots\ldots \quad (5)$$

when combined with eqns. (1), (2) and (3) enables a (P, ρ) point to be evaluated for each set of ρ_0, u, D data.

A more exact relation between u and σ has been given by Goranson *et al.* (1955, eqn. (10)), but the thermodynamic data necessary to evaluate this equation are not available for Perspex. However, Walsh and Christian (1955) have given relations setting upper and lower limits to the ratio σ/u. These are

$$\left(\frac{\sigma}{u}\right)_{\max} = \left[\left(\frac{P+2\beta}{2P}\right)\log_e\left(\frac{P+\beta}{\beta}\right)\right]^{1/2}, \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where $\beta = C_p\rho_0/\alpha$, and C_p is the specific heat at constant pressure and α the thermal coefficient of volume expansion of Perspex, and

$$\left(\frac{\sigma}{u}\right)_{\min} = \frac{\int_0^P \left[\left(-\frac{\partial v}{\partial p}\right)^{1/2} dp\right]_H}{[P(v_0 - v)]^{1/2}} \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where the subscript H denotes integration along the pressure–volume curve. An approximate pressure–volume curve can be derived using eqn. (5) and used to evaluate eqn. (7). This can then be used to construct a corrected pressure–volume curve and the process repeated until self-consistency is obtained. This method was not employed since any error introduced by the approximation $u = \sigma$ is less than the inherent experimental errors (see § 5.3).

§ 3. Experimental Method

3.1. *Aim*

The experiments were designed to measure the velocity of propagation of a shock wave passing through Perspex, and also the velocity of the free surface caused by the normal reflection of the shock wave at the Perspex–air boundary. The shock wave was produced by detonating an explosive charge in contact with Perspex, the maximum pressure generated in the Perspex being of the order of 10^{11} dynes/cm^2. The Perspex was in the form of blocks of various thicknesses.

3.2. *Explosive*

The explosive charge consisted of a mild steel tube measuring 3 in. long by 3 in. internal diameter, with $\frac{1}{4}$ in. walls, filled with plastic explosive (P.E. No. 2) (see fig. 1 (*a*)). Initiation was achieved by means of a detonator, electric, No. 8 Briska, placed in contact with a tetryl pellet, $\frac{1}{2}$ in. dia. by $\frac{1}{4}$ in. long, which was countersunk flush into the top surface of a detonation wave-shaper. The latter generated a plane detonation wave in the plastic explosive charge.

3.3. *Target Plates*

The test plates were 4 in. × 4 in. and cut from Perspex sheet up to thicknesses of 1 in. Greater thicknesses were built up by carefully placing sheets together and taping with Sellotape. The two large surfaces were

machined to be parallel and were flat to within 0·0015 in. over a diagonal. The average roughness was about 15 micro-inches.

3.4. *Probes*

Three types of double probe were used in the experiments to determine shock and free surface velocities. The probes acted as high speed switches in the event circuit shown in fig. 1 (*c*) and consisted essentially of two poles, at a potential difference of 300 volts, separated by about 0·002 in. The poles were joined electrically when a shock front produced intense ionization across them, thus discharging the event circuit condenser into a matched coaxial cable which transmitted a pulse directly to the deflector plates of a high speed oscilloscope.

A probe made from two nylon-covered copper wires (32 s.w.g.), glued together as a flat twin arrangement, was placed in the P.E. surface where it met the Perspex surface. The arrival of the detonation front at the probe caused the two wires to short together and thus operate the event circuit. The twin-wire probe used for the shock velocity measurements was made from two nylon-covered copper wires (32 s.w.g.) twisted together. The probe end of the wires was cut level and carefully rubbed flat on abrasive paper. The insulation of the probe was then checked at 500 volts.

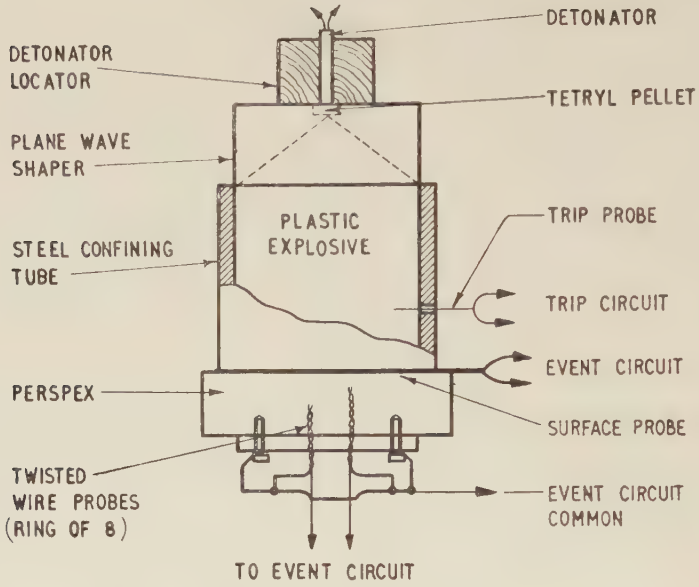
For the determination of the free surface velocity, a probe of more rigid construction than the twisted wire type, but still of small end-area, was required. A double probe was developed of concentric form which consisted of a copper tube at earth potential, 0·02 in. bore and 0·032 in. outside diameter, containing a nylon-covered wire glued in the bore so that the base of the wire was inset 0·002 in. with respect to the end of the tube : the end of the tube and wire had been carefully rubbed flat. The probe was soldered into a hole drilled along the axis of a 6 B.A. brass screw. This enabled the probe to be easily set in its carrier to the required distance above the free surface. A small piece of Sellotape covered the end of the probe thus enclosing a bubble of air which would ionize readily when compressed. This was necessary as, just before firing the charge, the probe was enclosed in an atmosphere of propane which does not ionize easily.

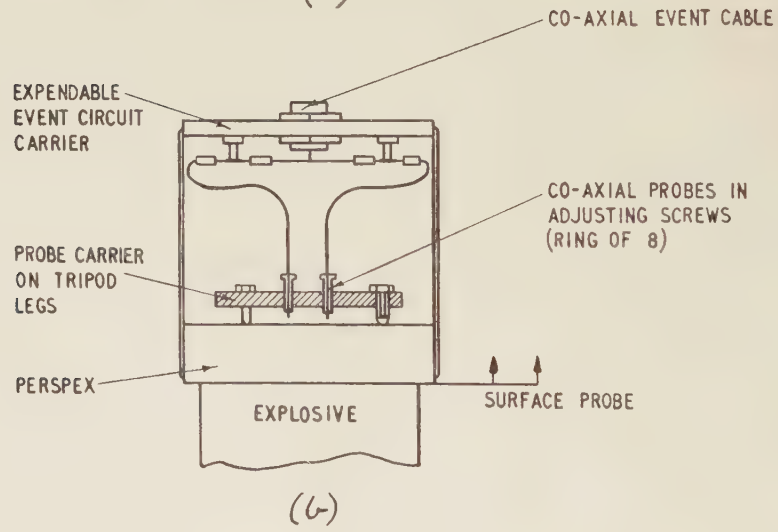
An account of experiments with these and other forms of probe is given in § 5.2.

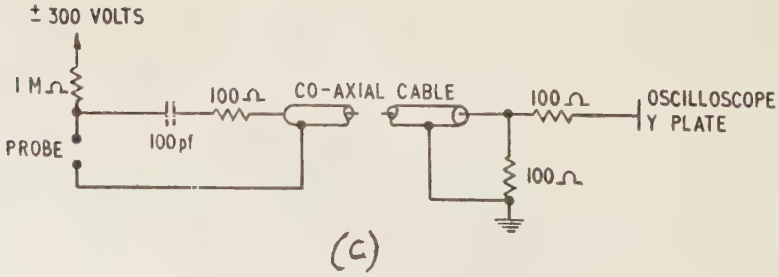
3.5. *Shock Wave Velocity*

The shock wave velocity was determined by measuring the time of arrival of the shock front at measured distances apart (fig. 1 (*a*)). For these experiments, the explosive was placed vertically above the test plate. Eight twisted wire probes were placed into flat-bottomed holes drilled on a circle of 0·6 in. dia. at the centre of a 1 in. thick Perspex plate. The holes differed in length by $\frac{1}{8}$ in. and were grouped so that

Fig. 1







Experimental arrangements.
(*a*) Shock wave velocity measurement.
(*b*) Free surface velocity measurement.
(*c*) Basic event circuit.

distortion of the shock wave produced when the shock reached a probe did not affect those parts of the wave which would encounter the following probes. In the first series, the probes were in contact with the bases of the holes, and in the second they were set up 0·002 in. from the bottom of the holes. Each probe was connected into the standard event system (fig. 1 (*c*)), and functioned as described in § 3.4 above.

3.6. *Free Surface Velocity*

The free surface velocity was determined by placing concentric probes at known distances above the surface (0·005 in. to 0·110 in.), fig. 1 (*b*), and obtaining from oscilloscope records the time of arrival of the surface at these probes. Eight of the probe screw assemblies described in § 3.4 were fitted into a brass plate with tripod legs, and this rested on the test surface. The probes were equi-spaced along the circumference of a circle of 0·6 in. diam. at the centre of the Perspex. Since the probe carrier rested on the test surface under its own weight, it was necessary to place the charge vertically below the target plate. The probe carrier served as a common earth and the centre wire of each probe was connected to an expendable event system mounted above the probes.

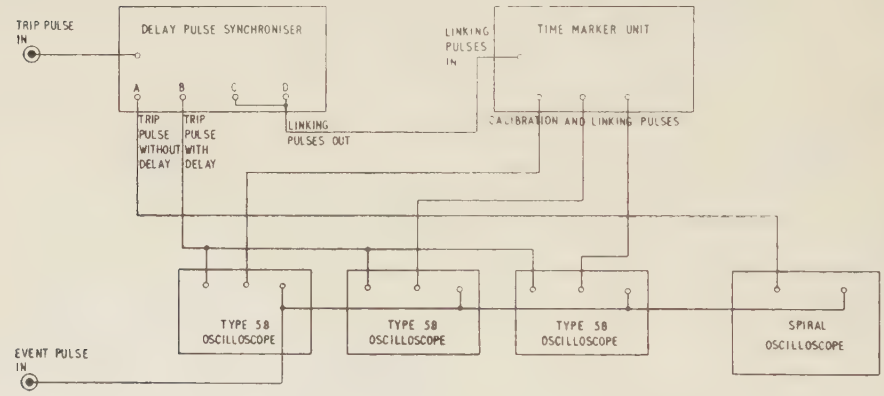
3.7. *Recording*

Three high velocity linear oscilloscopes and one spiral oscilloscope were used to record the time between event pulses. Initially Mark IIB linear oscilloscopes were used, which had been designed in the A.R.E. Instrument Section, but subsequently improved versions, Type 58, developed by Mr. Barnsley, were used. The Type 58 oscilloscope contains a new cathode-ray tube, 4EP1, which gives smaller spot size, greater deflection sensitivity and improved frequency response in comparison with the cathode-ray tube used in the Mark IIB. No published figures for the performance of the two types of oscilloscope are available. The following are approximate values under the conditions of the experiment, Type 58 oscilloscope quoted first : deflection sensitivity, 60 volts/cm ; 120–140 volts/cm : frequency response, 400 Mc/s ; 200 Mc/s : spot size, 0·55 mm ; 0·9 mm. The velocity of the spot across the cathode-ray tube may be varied so that the transit time will be 0·1 or 0·2 or 0·5 increasing by a factor of 2^n to 128 μsec. The Type 58 oscilloscope functions adequately on all these ranges whereas the Mark IIB oscilloscope is not satisfactory on sweep times of less than 2 μsec duration. For the experiments on Perspex, the time ranges used were of 0·5, 1 or 2 μsec duration, depending on the total time that event pulses were expected to occupy. The spiral oscilloscope was of the blanking type with a collapsing spiral, one revolution of the spot being crystal-controlled to take 2 μsec. The event pulses were fed into the cathode of the cathode-ray tube, positive pulses causing gaps in the trace and negative ones

brightening the trace. A one shot camera containing a Wray $f1{\cdot}0$ lens, giving a reduction in image-size of 4 : 1, was fitted to each oscilloscope to photograph the trace. The time resolution attained was ± 5 mμsec on a microsecond trace in the case of the Type 58 oscilloscopes, and ± 10 mμsec for the spiral oscilloscope.

A block diagram, fig. 2, outlines the electronic equipment used and the method of connection. Coaxial cable, of 100 ohms impedance, was used for all connections in the pulse circuits. A twisted wire probe inserted into the charge triggered the D.P.S. unit (Delay Pulse Synchronizer) about 3 μsec before the detonation wave would have reached the Perspex surface (fig. 1 (*a*)). The function of the D.P.S. was to emit positive pulses, either without delay (Channel A) or after selected delays (Channels B, C and D) with respect to the trigger pulse. The spiral oscilloscope was tripped from Channel A of the D.P.S. and recorded the complete sequence of event pulses, all of which occurred within 30 μsec.

Fig. 2



Block diagram of recording system.

The linear oscilloscopes were triggered from Channel B after a set delay which ensured that all the event pulses with the exception of that due to the surface probe would be displayed. The linear oscilloscopes were connected in cascade, the start of the traces on the second and third being delayed with respect to the first and second, respectively, using the beam delay control on the oscilloscope. A crystal-controlled time-marker unit fed pulses on to one Y-plate of each oscilloscope for the purpose of calibrating the traces : the other Y-plate received the event pulses. The three traces were linked by displaying two marker pulses, each one being common to consecutive pairs of traces with the minimum of trace overlap. These pulses were obtained from Channels C and D of the D.P.S. via the time-marker unit. Thus the time resolution was

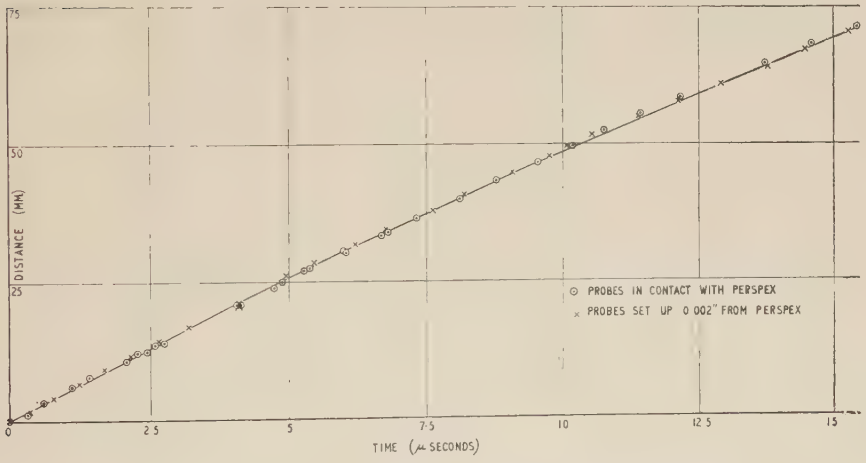
increased in comparison with the simpler arrangement employing one oscilloscope. The spiral oscilloscope, in addition to recording the surface probe pulse, enabled a satisfactory result to be obtained if the linear oscilloscopes failed, due to faulty delay setting, to record the event pulses.

The oscilloscope traces were photographed on Ilford H.P.3 and developed in I.D.33. Typical records are shown in fig. 3, Pl. 102.

§ 4. Experimental Results

The results of experiments on shock velocity measurement are shown in graphical form, fig. 4. The common zero for each experiment, irrespective of block thickness, was the tripping of the surface probe. This was assumed to trigger immediately the detonation wave reached it, so that a correction was applied for the time taken for the shock to reach the Perspex. No difference was observed between the results obtained

Fig. 4



Distance versus time plot of shock front in Perspex.

from probes in contact with Perspex and those set up 0·002 in. In the latter case, a correction—varying from 0·01 to 0·03 μsec—was made to the times recorded. This correction represented the time taken by the air shock to travel from the measured end of the hole to the probe.

A quadratic equation, calculated by the method of least squares, was found to fit the exprimental points :

$$t = 0{\cdot}183s + 0{\cdot}0005s^2 \qquad . \; . \; . \; . \; . \; . \; . \; (8)$$

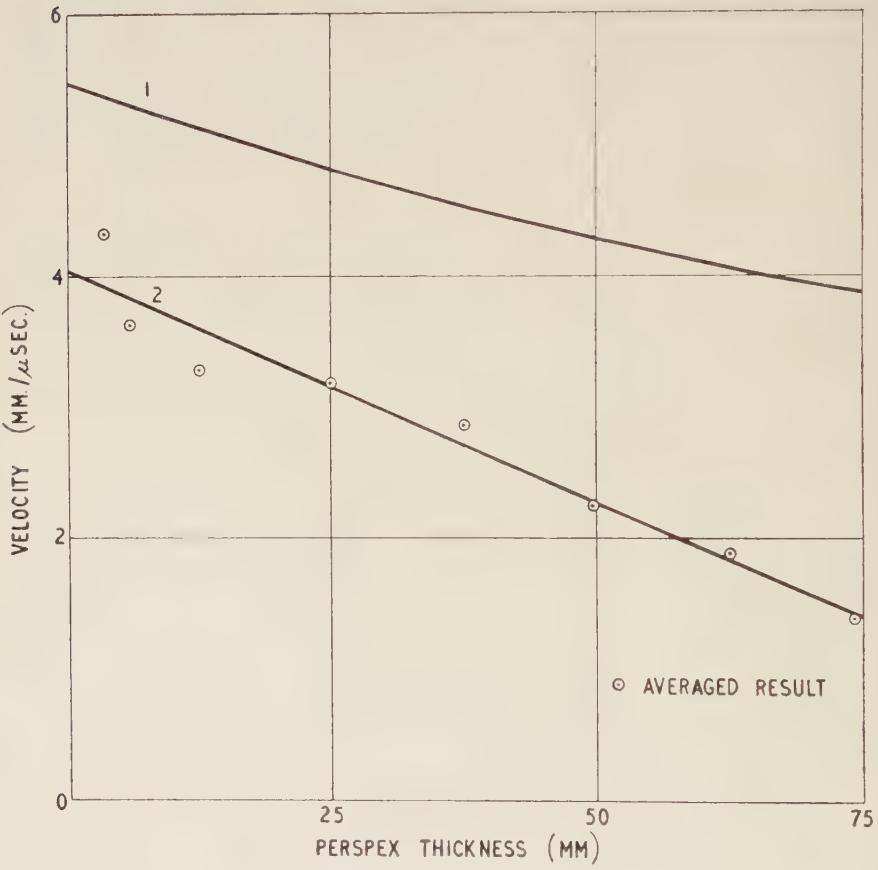
where t is in μsec, s in millimetres. The velocity at a point s within

the range covered by the experiments ($s \leqslant 75$ mm) was obtained by differentiating eqn. (8) :

$$\text{shock velocity, } D = \frac{1}{0{\cdot}183 + 0{\cdot}001s} \quad . \quad . \quad . \quad . \quad (9)$$

where D is in mm/μsec. The equation gives a maximum shock velocity of 5·47 mm/μsec. Figure 5, curve 1, shows the variation of shock wave velocity with Perspex thickness over the range covered by the experiments.

Fig. 5



Curve 1 : Shock wave velocity—calculated from fig. 4.
Curve 2 : Free surface velocity—calculated from table 1.
Shock wave velocity and free surface velocity versus Perspex thickness.

Table 1 gives the results for the free surface velocity experiments on thicknesses ranging from $\frac{1}{8}$ in. to 3 in. The average velocity results were plotted for each thickness, and a best fitting straight line calculated.

This is shown in fig. 5, curve 2. The equation of the line is :

$$\text{free surface velocity, } V = 4{\cdot}03 - 0{\cdot}035s. \quad . \quad . \quad . \quad (10)$$

Using eqns. (1), (2), (3), (5), (9) and (10), the dynamic compression (Hugoniot) curve, fig. 6, was constructed. A value of 1·2 g/cm^3 was used for the density of Perspex.

Table 1. Perspex : Results of Free Surface Velocity Experiments

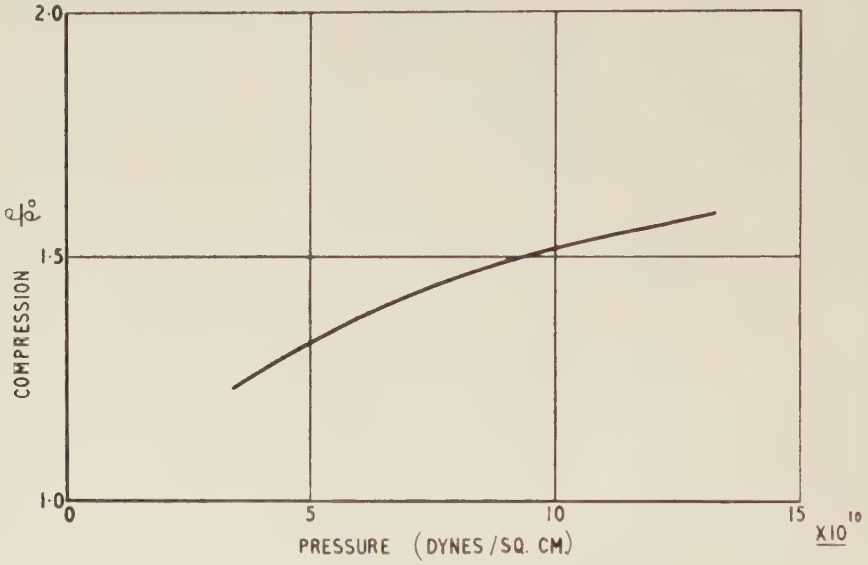
Perspex thickness		Free surface velocity (mm/μsec)	Standard deviation	Mean	Standard deviation of mean
(Nominal) in.	(Actual) mm				
$\frac{1}{8}$	3·18	3·92	±0·09	4·33	—
	3·27	4·06	±0·16		
	3·20	5·03	±0·28		
$\frac{1}{4}$	5·97	3·85	±0·25	3·63	±0·13
	6·17	3·28	±0·11		
	6·22	3·95	±0·20		
	6·22	3·67	±0·29		
	6·22	3·38	±0·12		
$\frac{1}{2}$	12·62	3·31	±0·18	3·28	±0·10
	12·39	3·34	±0·06		
	12·22	3·55	±0·06		
	12·55	2·92	±0·11		
	12·34	3·28	±0·07		
1	24·93	2·98	±0·06	3·19	—
	24·93	3·05	±0·06		
	24·77	3·53	±0·08		
$1\frac{1}{2}$	37·5	3·08	±0·05	2·87	—
	37·4	2·66	±0·18		
2	49·6	2·24	±0·06	2·24	—
	49·7	2·23	±0·06		
$2\frac{1}{2}$	61·6	1·98	±0·13	1·89	—
	63·1	1·80	±0·02		
3	75·0	1·39	±0·02	1·40	—
	73·8	1·41	±0·06		

After detonating the explosive charge in contact with the Perspex plates, a few small fragments of Perspex were recovered in the case of the 3 in. thick target plates only. An odour, characteristic of vaporized Perspex, could be detected in the firing chamber.

§ 5. Errors

The errors inherent in these experiments are, in the main, random and are discussed below together with a possible more serious systematic error due to the gaseous shock which precedes, and is generated by, the moving free surface.

Fig. 6



Calculated from eqns. (1), (2), (3), (5), (9) and (10).
Dynamic compression curve for Perspex.

5.1. *Random Errors*

It is implicitly assumed in the measurements of both shock and free surface velocity that the shock wave refracted into the Perspex block is plane and parallel to the surfaces of the block. Any deviation from this condition would have resulted in some of the probes being struck earlier or later than should have been the case. Since the probes were arranged to increase in distance in both directions from the first probe, the usual effect would have been an increased scatter in the times recorded. A study of the results obtained indicates that the error from the skewness of, and irregularities in, the shock wave was not great.

The shock velocity results for blocks of different thicknesses were combined by taking the probe event at the explosive Perspex interface as the time zero. Due to variations in the surface flatness of the Perspex block, this probe could have varied in its distance from the other probes by $\pm 0{\cdot}0015$ in. ($\pm 0{\cdot}04$ mm). This is equivalent to a time error of $\pm 0{\cdot}007$ μsec at the maximum velocity and since the times were measured to $\pm 0{\cdot}01$ μsec the resultant error in the time zero should be less than $0{\cdot}02$ μsec. The twisted wire probes inside the Perspex block were set in holes whose depths were measured to an accuracy of $\pm 0{\cdot}0002$ in.

($\pm 0{\cdot}005$ mm). The position of these probes might have varied due to the way the probe was pushed into the hole and glued in position, or handling and temperature effects might have caused movements in the probes after setting. The error in the shock velocity determination was estimated from the combined distance-time results (fig. 4) as less than $\pm 2\%$ and the major source of error was in the combining of the results from different experiments.

In the case of the free surface velocity the probes were set in the probe carrier to the required distance on a flat table fitted with a dial gauge. This enabled probe distances to be measured to $\pm 0{\cdot}0001$ in. The probe carrier was then transferred to the Perspex surface which was not as flat as the measuring table. The error in the distance from this operation was minimized by arranging that the probes were set on a circle at the centre of the test plate. The design was such that temperature effects were self-cancelling but handling might have caused some variation. Another possible source of error was in the probe. The centre wire might not have been exactly 0·002 in. inset and this would give an error in the time of an event. The centre wire is unlikely to have been misplaced by more than $\pm 0{\cdot}002$ in. and the corresponding variation in time is $\pm 0{\cdot}01$ μsec. All the above errors would lead to increased scatter of the experimental points about the ' true ' line.

The experimental values obtained for the free surface velocity are given in table 1. The accuracy is generally better than $\pm 5\%$ for each experiment though the variation between nominally identical rounds is much larger in some cases. (This is discussed later.) The scatter of the experimental points is on the average only slightly larger than that expected from the accuracy of reading the oscilloscope traces ($\pm 0{\cdot}005$ μsec corresponding to $\pm 0{\cdot}0006$ in. in distance) indicating that the effect of all the other sources of error is about the same size as the time resolution.

5.2. *Systematic Errors*

The movement of the free surface sets up a shock wave in the gas ahead of the surface. This shock wave moves with a greater velocity than the surface and, if the probe system responds to the shock, an enhanced velocity will be recorded. The relation between the free surface velocity and the shock velocity in air and in propane has been calculated (Deas, private communication).

Various experiments with probes of differing anticipated sensitivity to shock were carried out, and the results are given in table 2. Four types of probes were used : a brass screw making contact with a conducting surface, a concentric probe with Sellotape over the end, a concentric probe without Sellotape and a twisted wire probe. The results using $\frac{1}{8}$ in. thick aluminium plate indicate the relative merits of the systems. The brass screw array in propane gives the true free surface velocity, while the concentric probes with Sellotape in propane or in air give mean results which are not significantly different (one result is high and would

agree with the propane velocity). However, the concentric probes without Sellotape in air give a significantly higher result which corresponds to the expected air shock velocity.

The results using the 2 in. thick Perspex show that the free surface velocity is approximately the same as that of the $\frac{1}{8}$ in. thick aluminium sheet, and thus we would expect the concentric probe with Sellotape to give the correct velocity. However, there is a tendency to get a higher result when the probes are surrounded by air rather than propane.

Table 2. Measurement of Free Surface Velocity. Experimental Results from Various Probe Systems in Air and Propane

Test material	Gas	Probe system	Velocity (mm/μsec)
Aluminium $\frac{1}{8}$ in. thick	Propane	Brass screw	$2{\cdot}34 \pm 0{\cdot}06$
	do.	Concentric, Sellotape	$2{\cdot}32 \pm 0{\cdot}12$ $2{\cdot}56 \pm 0{\cdot}07$
	do.	Brass screw † Concentric, Sellotape †	$2{\cdot}20 \pm 0{\cdot}10$ † $2{\cdot}28 \pm 0{\cdot}07$ †
	Air	Concentric, Sellotape	$2{\cdot}36 \pm 0{\cdot}07$
	do.	Concentric, *no* Sellotape	$2{\cdot}73 \pm 0{\cdot}07$
Perspex 2 in. thick	Propane	Concentric, Sellotape	$2{\cdot}24 \pm 0{\cdot}06$ $2{\cdot}23 \pm 0{\cdot}06$ $1{\cdot}95 \pm 0{\cdot}14$
	Air	Concentric, Sellotape	$2{\cdot}37 \pm 0{\cdot}09$ $2{\cdot}23 \pm 0{\cdot}06$ $2{\cdot}47 \pm 0{\cdot}10$
	do.	Concentric, *no* Sellotape	$2{\cdot}47 \pm 0{\cdot}05$
Perspex $\frac{1}{4}$ in. thick	Propane	Concentric, Sellotape	$3{\cdot}85 \pm 0{\cdot}25$ $3{\cdot}28 \pm 0{\cdot}11$ $3{\cdot}95 \pm 0{\cdot}20$ $3{\cdot}67 \pm 0{\cdot}29$ $3{\cdot}38 \pm 0{\cdot}12$
	do.	Twisted twin wire, *no* Sellotape	$3{\cdot}83 \pm 0{\cdot}08$
	Air	Concentric, Sellotape	$4{\cdot}10 \pm 0{\cdot}10$
	do.	Concentric, *no* Sellotape	$4{\cdot}07 \pm 0{\cdot}03$
Perspex 1 in. thick	Propane	Twisted twin wire, *no* Sellotape	No events. Insufficient ionization to trigger probes

† Probes mixed, four of each type, same target plate.

The third group of firings using $\frac{1}{4}$ in. thick Perspex gave a large variation in the results from experiments of the same type. One interpretation is that in some cases we were recording the movement of the shock front and in others the true free surface velocity. The

value obtained for the air shock velocity using the concentric probes without Sellotape supports this interpretation, as does the high velocity figure obtained in propane using an uncovered twisted wire double probe. However, the experimental errors in these firings are of the same magnitude as the difference between propane shock and surface velocity and a definite decision must await improvement of the probe system. Any error introduced by this will only affect the high pressure end of the pressure–volume curve and will be no larger than that due to the scatter in results. To keep this uncertainty as small as possible the probes were immersed in propane rather than air, since there is a larger difference between the free surface velocity and the resulting air shock velocity than between the free surface velocity and the resulting shock velocity in propane.

5.3. *Observations*

In order to obtain an insight into the effect on the calculated Hugoniot of the experimental errors in the shock velocity and free surface velocity results, the following calculations were made. A nominal thickness of Perspex was considered and the shock velocity calculated from eqn. (9) increased or decreased by 2%. These values were then combined with the maximum and minimum values for the free surface velocity at the selected thickness (from table 1) in such a way as to give the greatest range in values for the compression ratio. The results obtained at various thicknesses are given in table 3 together with the values for the pressure at the corresponding thicknesses calculated from the shock velocity (eqn. (9)) and the mean free surface velocity.

Table 3

Pressure (dynes/cm)	Compression ratio range (ρ/ρ_0)
$4{\cdot}5\times10^{10}$	1·28 to 1·33
$9{\cdot}0\times10^{10}$	1·44 to 1·60
$12{\cdot}0\times10^{10}$	1·46 to 1·66
$12{\cdot}6\times10^{10}$	1·56 to 1·96

The particle velocity was taken as half the free surface velocity. The possible maximum and minimum values for the particle velocity at various pressures were calculated using eqns. (6) and (7). The specific heat of Perspex at constant pressure was found to be 0·33 cal g^{-1} deg c^{-1}. The compression ratio obtained from these values varied from 1·47 to 1·52 at a pressure of 10^{11} dynes per cm^2, the discrepancy decreasing rapidly as the pressure reduced. The error introduced by taking $u = \frac{1}{2}$ free surface velocity may thus be neglected in comparison with the experimental errors.

§ 6. Discussion

These experiments have enabled compression data for Perspex to be obtained in a region not covered by other experiments. The results can be fitted by an expression of the form

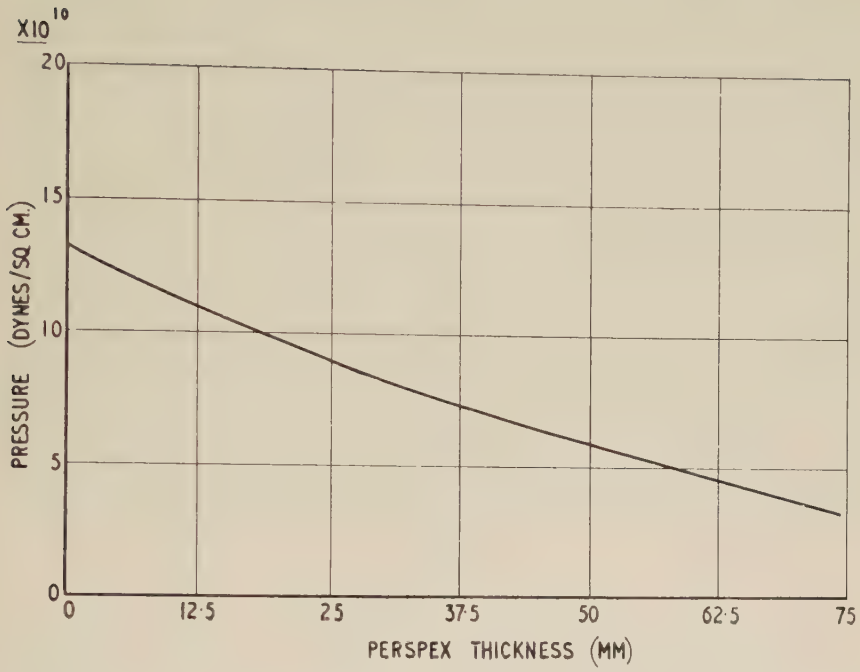
$$P = A\left(\frac{\rho-\rho_0}{\rho_0}\right) + B\left(\frac{\rho-\rho_0}{\rho_0}\right)^2 \quad . \quad . \quad . \quad . \quad . \quad (11)$$

where $A = 6{\cdot}64 \times 10^{10}$ dynes/cm^2 and $B = 26{\cdot}2 \times 10^{10}$ dynes/cm^2. This fits the results well over the range covered by the experiment but extrapolation to greater pressures must be used with great care. The only other published information in this region is the single point which can be obtained from the results of Lawton and Skidmore (1956). Their point ($P = 1{\cdot}77 \times 10^{11}$ dynes/cm^2, $\rho/\rho_0 = 1{\cdot}63$) agrees well with the values of $\rho/\rho_0 = 1{\cdot}71$ given by eqn. (11) for the same pressure. Pressures up to $\sim 10 \times 10^{10}$ dynes/cm^2 have been reached statically but no results have been published for Perspex under these conditions. It would be very interesting to be able to compare the values of A and B obtained in a slow isothermal experiment with the values obtained in the present experiment.

It is possible to compute the velocity of propagation of elastic dilatational waves from eqn. (11) (i.e. as $P \to 0$) but this involves an extrapolation of the experimental results and no great reliance can be placed on the value obtained. The velocity thus obtained is 2·4 mm/μsec. Kolsky (1956) has published results which enable a velocity of 2·3 mm/μsec to be deduced for the velocity of longitudinal waves along rods of Perspex. Kolsky, private communication, suggests that this value of the longitudinal velocity corresponds to a dilatational wave velocity of 2·5–2·6 mm/μsec, which agrees reasonably well with the prediction of eqn. (11). Equation (9) is an empirical relation between shock velocity and thickness of Perspex and is valid only over the range covered by experiment. It is obviously incorrect at large thicknesses of Perspex as it indicates a continuing decrease in shock velocity with increasing Perspex thickness instead of an asymptotic approach to the sonic velocity as given by eqn. (11).

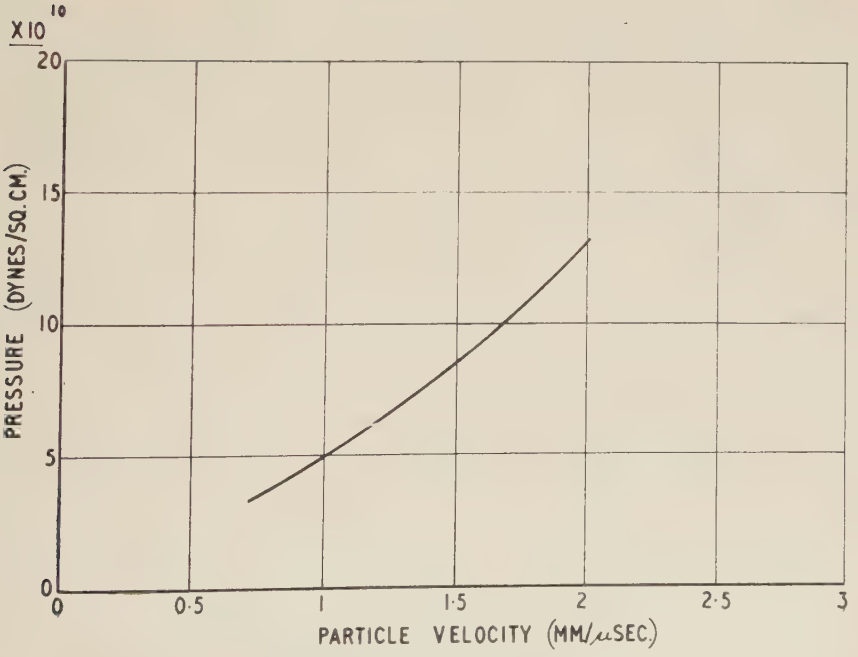
A feature of the experimental results for the velocity of the free surface (tables 1 and 2) is that there was a greater round to round variation than might be expected from the standard deviation of each experiment. The source of this round to round variation is not clear ; it could be due to variability in the explosive charge, in the refraction of the shock wave into the Perspex or in the Perspex or a combination of all three. The effect of the first two could be eliminated if it had been possible to measure both shock velocity and free surface velocity in the one experiment, but this was not possible due to the limitation on the size of explosive charge which could be used. The other possible source of the round to round variation is that in some of the free-surface velocity experiments some of the probes were responding to the shock in the

Fig. 7



Pressure versus Perspex thickness.

Fig. 8



Pressure versus particle velocity.

propane gas as mentioned in § 5.2. This interpretation is supported by the results which show that there is a smaller round to round variation in the shock velocity experiments where this effect is absent. To obtain greater accuracy in the experimental results it will be necessary to improve the probe system used to detect the movement of the free surface, particularly at the higher velocities. It is believed that with some development this difficulty should be overcome and more consistent results obtained.

In addition to the graph of P *vs* ρ/ρ_0 the results have been plotted in the form of pressure against thickness of Perspex (fig. 7) which gives the effective attenuation of a plane wave in Perspex. This is useful in the design of systems where high amplitude shock loading occurs. The slope of the graph of P *vs* u (fig. 8) is the shock impedance (ρD) of Perspex which is seen to increase steadily with increasing pressure. This means that the shape of a high amplitude stress pulse will be changed when it is refracted at a Perspex boundary since the amplitude of the refracted wave depends on ρD. Figure 8 will also be useful in calculations on high speed impact where pressures of this order are met.

Acknowledgments

The authors would like to thank Mr. D. A. Barnsley who was responsible for the design of the Type 58 oscilloscopes and Mr. H. G. Peppiatt who assisted in the experimental work. Reproduced with the permission of the Controller, H.M.S.O.

References

Bancroft, D., Peterson, E. L., and Minshall, S., 1956, *J. appl. Phys.*, **27**, 291.
Bridgman, P. W., 1945 a, *Proc. Amer. Acad. Arts Sci.*, **74**, 425 ; 1945 b, *Ibid.*, **76**, 9 ; 1945 c, *Ibid.*, **76**, 55 ; 1949, *Ibid.*, **77**, 189.
Deal, W. E., 1957, *J. appl. Phys.*, **28**, 782.
Duvall, G. E., and Zwolinski, B. J., 1955, *J. acoust. Soc. Amer.*, **27**, 1054.
Goranson, R. W., Bancroft, D., Burton, B. L., Blecher, T., Houston, E. E., Gittings, E. F., and Landeon, S. A., 1955, *J. appl. Phys.*, **26**, 1472.
Kolsky, H., 1956, *Phil. Mag.*, **1**, 693.
Lawton, H., and Skidmore, I. C., 1956, *The Physical Chemistry of Processes at High Pressure* (London : The Faraday Society), p. 188.
Minshall, S., 1955, *J. appl. Phys.*, **26**, 463.
Pack, D. C., Evans, W. M., and James, H. J., 1948, *Proc. phys. Soc. Lond.*, **60**, 1.
Rice, M. H., and Walsh, J. M., 1957, *J. chem. Phys.*, **26**, 824.
Walsh, J. M., and Christian, R. H., 1955, *Phys. Rev.*, **97**, 1544.
Walsh, J. M., and Rice, M. H., 1957, *J. chem. Phys.*, **26**, 815.
Walsh, J. M., Rice, M. H., McQueen, R. G., and Yarger, F. L., 1957, *Phys. Rev.*, **108**, 196.

Dielectric Properties of Some Metaniobate and Metatantalate Ceramics†

By R. V. Coates‡ and H. F. Kay
H. H. Wills Physics Laboratory, University of Bristol

[Received August 15, 1957]

Abstract

The preparation and basic dielectric properties of the metaniobates and metatantalates of iron, manganese, lead, tin, calcium, strontium and barium (all in the divalent state), and trivalent aluminium are described. Manganese metatantalate, barium metaniobate, and aluminium metaniobate have been found to be ferroelectric, and lead metatantalate antiferroelectric. These are in addition to ferroelectric lead metaniobate discovered by Goodman in 1953. Strontium and barium metatantalates may have transition points below $-183°$c.

§ 1. Introduction

Surveys of the properties of perovskite type materials of general formula $A^n B^{6-n} O_3^{-2}$, together with a review of the whole field of ferroelectrics, can be found in works by Shirane, *et al.* (1955) and Känzig (1957). In 1953 Goodman prepared lead metaniobate, $PbNb_2O_6$, and found it to be ferroelectric. If the general perovskite formula for $n = 1$ is written as $A^{1+}(BO_3)^{1-}$, then the formula which includes lead metaniobate is $A^{2+}(BO_3)^{1-}{}_2$, which can be regarded as the second example in a sequence where the valency of the A ion increases by unity each time, and B remains an ion from group 5.

Metaniobates and tantalates of divalent metals occur naturally as the two dimorphous mineral groups columbite and tapiolite, which are each finely graded series having the general composition iron-manganese metaniobate-metatantalate. Many specimens of these minerals were obtained from the Bristol Museum, British Museum, and Universität Museum of Oslo, to whom the authors are indebted for their cooperation. However, despite good or excellent habit the crystals were mesomorphic, and the axes of the constituent crystallites were poorly defined; some had one axis common to all crystallites, while others had no common axis at all. Moreover, the electrical conductivity was so high as to preclude any reliable electrical information being obtained so that all specimens had to be produced in the laboratory.

The research to be described surveys the dielectric properties of the metaniobates and metatantalates of iron, magnanese, lead, tin, calcium,

† Communicated by the Authors.
‡ Now at Fisons Ltd., Levington Research Station.

strontium and barium, all in the divalent state. In order to make a preliminary investigation into the third formula type, $A^{3+}(BO_3)^{1-}{}_3$, the metaniobate and tantalate of aluminium were also prepared.

As ceramic specimens only were investigated, and no single crystal electrical measurements made, the results are not presented as a complete description of the electrical properties of these materials, but rather for the guidance of those concerned in dielectrics research.

§ 2. Specimen Preparation

The starting materials were the pentoxide of niobium or tantalum and the sulphate, carbonate, or oxide of the A metal. Whenever possible compounds of the A metal were used which decompose at a high temperature to give the required oxide, for these are far more reactive (e.g. Cohn 1948, Goodman 1953). In most cases alternative starting materials were tried for the same final product to ensure that the desired form was obtained.

Approximately stoichiometric amounts of the two starting materials were ground together under acetone, and, when almost dry, pressed into discs of 0·5 in. diameter by 0·2 in. thick. The prefiring temperatures were about 50°C above the decomposition temperature of the sulphate or carbonate used, and this was maintained for 2 hours, achieving the correct value from 400°C in $1\frac{1}{2}$ hr.

After regrinding under acetone and drying off to a putty-like consistency the final discs were pressed at 10 tons/sq. in. in a stainless steel die, and fired for 24 hr. at about 50°C below the fusion point, provided this did not exceed 1500°C. All firing was performed with an oxygen atmosphere which, together with a 5% excess of the B_2O_5 constituent in the original mass, gave specimens of lowest conductivity. X-ray powder diffraction photographs were taken at each stage to check and follow production of the required material.

The disc faces were ground plane parallel, the disc porosity measured by water absorption, and finally aluminium electrodes were evaporated on to the faces.

§ 3. Methods of Measurement

For all electrical measurements the disc was held in a combined specimen holder and heating muffle. Below room temperature measurements were made at two fixed points—the temperature of boiling liquid oxygen, and that of vaporising carbon dioxide. Here, the disc was first maintained at 150°C for half an hour to ensure its being thoroughly dry, and then plunged rapidly beneath the coolant. Sufficient time was allowed for the assembly to reach the required temperature before any measurements were taken.

Dielectric constants were measured in two ways. First, a modified Sawyer Tower (1930) display apparatus was used to give a visual indication of the displacement, (D), to field, (E), relation at 50 c/s. An indication of

the presence of ferroelectricity is given by saturation of the displacement at the tips of the loop. The peak to peak dielectric constant at the given applied field can be calculated from the geometry of the loop and from a knowledge of the disc electrode area and disc thickness. The variation of dielectric constant with temperature for a fixed field value, and the variation of dielectric constant with applied field at constant temperature were measured for each material investigated.

The dielectric constant values so obtained are the moduli of the complex dielectric constants, and therefore each includes a contribution from the loss component. When the conductivity of the specimen becomes too large, generally for temperatures above about 200–250°C, the real component of the modulus is swamped, and an inaccurate value of dielectric constant results.

The second method of dielectric constant measurement was by means of a Q-meter. This gives the real part of the complex quantity only, and therefore in general will be smaller than the value obtained on the display apparatus. Moreover, since fields of only a few volts/cm instead of many kv/cm are used, the incremental value is obtained which is again usually smaller than the peak to peak value for non-linear dielectric materials. All discs were measured for dielectric constant and loss variation with temperature up to 500°C, at frequencies of 0·1, 1 and 5 Mc/s.

All quoted values of dielectric constant have been corrected for specimen porosity by the formula due to Van Santen (1946):

$$k_r/k_a = 1/(1-\tfrac{3}{2}p)$$

where k_r is the real dielectric constant for no voids,

k_a is the measured dielectric constant,

p is the specimen porosity.

§ 4. Survey of Results

The main results for each material are given in the accompanying table. Additional details are given below for the most interesting cases, and full details of $\tan\delta$ *vs* T, $1/k$ *vs* T and k *vs* E for each specimen tabulated may be found elsewhere (Coates 1956).

4.1. *Lead Metaniobate*, $PbNb_2O_6$

The ceramics prepared by the authors gave clear saturation trends on their D/E loops, with 50 c/s dielectric constant value of 26 at −183°C rising to 90 at 200°C. No transition temperature was observed, as this is at 570°C which is higher than the range examined. The room temperature value of k, 75, is half that quoted by Goodman, which can be explained from the existence of a non-ferroelectric rhombohedral polymorph (Francombe 1956) produced for firing temperatures below 1200°C. Since the disc investigated here was fired first below 1200°C and second at 1230°C it is a mixture of both forms, as is confirmed by

Summarized Table of Results

Material	Starting chemicals	Firing temps. °c.	Colour	Density	Room temp. ; 50 c/s approx. 30 kv/cm k	tan δ	Remarks
$FeNb_2O_6$	$FeCO_3$; Nb_2O_5 (FeO; Nb_2O_5)	(i) 1300 (ii) 1320	Dead black	4·46	270 (at 1 Mc/s)	Very conductive	$k=18\pm2$ at -183°c and 1 Mc s. Too conducting however for detailed investigation.
$FeTa_2O_6$	$FeCO_3$; Ta_2O_5 (FeO; Ta_2O_5)	(i) 1330 (ii) 1370	Dark nigger brown	7·17	95 (at 1 Mc/s)	Very conductive	$k=50\pm10$ at -183°c and 1 Mc s. Difficulties as above. Not ferroelectric above -183°c.
$MnNb_2O_6$	$MnCO_3$; Nb_2O_5 (MnO; Nb_2O_5)	(i) 1320 (ii) 1350	Mid. brown	4·95	38	0·02	k/T relation almost horizontal. Not ferroelectric above -183°c.
$MnTa_2O_6$	$MnCO_3$; Ta_2O_5 (MnO; Ta_2O_5)	(i) 1320 (ii) 1350	Grey–brown	7·35	Probably 750	Very conductive	Too conductive for display apparatus use other than at -183°c. At this temperature $k=48$ and slight saturation visible. For r.f. k/T slope is positive at least up to 400°c. Possibly weakly ferroelectric.
$PbNb_2O_6$	(PbO; Nb_2O_5) $PbSO_4$; Nb_2O_5	(i) 1185 (ii) 1230	(Pale yellow) White, faint blue cast	6·2	75	0·14	Ferroelectric ; transition temperature above 200°c. See Goodman (1953). Mixture of ferroelectric orthorhombic and non-ferroelectric rhombohedral forms. See text.
$PbTa_2O_6$	(PbO; Ta_2O_5) $PbSO_4$; Ta_2O_5	(i) 1360 (ii) 1420	Pale blue-green	7·95	860	0·05	Not ferroelectric but marked dielectric constant peak at 70°c. Possibly antiferroelectric. See text and fig. 1.
$SnNb_2O_6$	SnO; Nb_2O_5	(i) 1350 (ii) 1420	Pale brownish-grey	4·76	160 at 1 Mc s	0·05	Very conducting at 50 c s and room temperature. k/T relation slope positive up to at least 400°c but due probably to barrier layer mechanism.

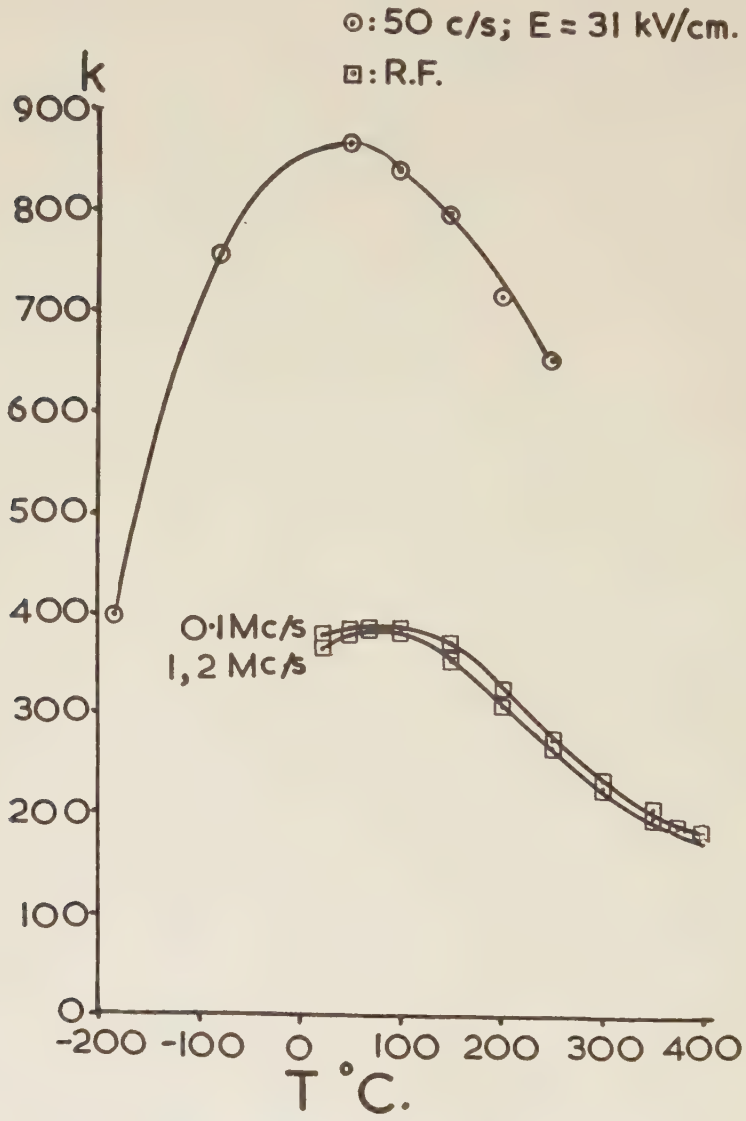
$SnTa_2O_6$	SnO; Ta_2O_5	(i) 1220 (ii) 1440	White to pale brown	7·06	100 at 1 Mc/s	0·06	As $SnNb_2O_6$. Investigated only up to 250°c.
$CaNb_2O_6$	(CaO; Nb_2O_5) $CaCO_3$; Nb_2O_5	(i) 1250 (ii) 1400	Pale cream	4·36	28	0·05	Not ferroelectric above −183°c. k/T slope almost horizontal.
$CaTa_2O_6$	(CaO; Ta_2O_5) $CaCO_3$; Nb_2O_5	(i) 1250 (ii) 1400	Pale cream	6·41	68	0·2	As for $CaNb_2O_6$. Room temperature value of k at 50 c/s too large due to large tan δ. At −183°c and 50 c/s $k=55$, tan $\delta=0{\cdot}03$.
$SrNb_2O_6$	(SrO; Nb_2O_5) $SrCO_3$; Nb_2O_5	(i) 1400 (ii) 1420	White	4·81	48	0·01	Not ferroelectric above −183°c. k *vs* T slope positive and increasing in magnitude at least up to 500°c.
$SrTa_2O_6$	(SrO; Ta_2O_5) $SrCO_3$; Ta_2O_5	(i) 1400 (ii) 1420	White	7·08	140	0·01	Not ferroelectric above −183°c. k *vs* T slope negative and increasing in magnitude from 0°c to −183°c. Possibly a transition below −183°c ; a plot of $1/k$ *vs* T is inconclusive.
$BaNb_2O_6$	BaO; Nb_2O_5	1220 (special conditions)	Fawn	5·44	130	0·05	Ferroelectric with transition temperature about 70°c. See text and fig. 2.
$BaTa_2O_6$	BaO; Ta_2O_5	(i) 1300 (ii) 1440	White	7·78	156	0·03	As $SrTa_2O_6$.
$AlNb_3O_9$	Al_2O_3; Nb_2O_5	(i) 1350 (ii) 1420	Pale cream	3·81	170	0·01	Ferroelectric with transition temperature at 220°c. See text and fig. 3.
$AlTa_3O_9$	Al_2O_3; Ta_2O_5	(i) 1350 (ii) 1420	White	6·13	65	0·02	Not ferroelectric above −183°c. k *vs* T slope level up to 400°c.

its x-ray powder diffraction pattern. The dielectric constant value is thus reduced by the amount of rhombohedral phase present.

4.2. *Lead Metatantalate*, $PbTa_2O_6$

Figure 1 shows the k *vs* T graph for this material. It is immediately obvious that some important transition occurs at about 70°c ; the curve of dielectric constant against temperature falls steeply above and

Fig. 1



Relation between dielectric constant and temperature at different frequencies for Pb_2TaO_6.

below this point for all frequencies. Despite this peak, the D/E trace at 50 c/s was a thin ellipse, tending to a straight line at -183°C, with no sign of any saturation effects. This suggests, therefore, that lead metatantalate is not ferroelectric, but might be anti-ferroelectric. If this is so, then at sufficiently high fields the trace should split into two small hysteresis loops, joined together by a linear portion where the field is insufficiently strong to cause a parallel arrangement of the dipoles. Cross (1954), working on sodium niobate, needed a field of 90 kv/cm to achieve this splitting, whereas the highest field applied in this instance was only 43·6 kv/cm. Similarly, Shirane and Pepinsky (1953) were unable to split the trace of lead hafnate with a field of 40 kv/cm, despite a dielectric constant peak of 540 at 215°C.

A plot of $1/k$ against T for the three sets of r.f. readings shows that above 70°C the points lie on three straight lines, confirming that this region obeys a Curie–Weiss law of the form $k=k_0+C/(T-\theta)$, with a fairly small contribution from k_0. From the straight lines, average values of the constants in this equation are : $k_0=4$; $C=85\,000$°K ; $\theta=-80$°C.

Lead metatantalate is probably anti-ferroelectric with a transition temperature at about 70°C.

4.3. *Barium Metaniobate*, $BaNb_2O_6$

The normal firing technique as described previously was not suitable for barium niobate : although it fuses at 1350°C, approximately, from about 1220°C upwards a ceramic disc slowly crystallizes into a matted mass of needle-shaped crystals, giving a very porous and fragile specimen quite unsuitable for investigation. The discs were therefore fired at only 1200°C, and for 48 hours at that temperature since the normal 24 hours was insufficiently long to give good sintering. The growth of microcrystals at 1200°C was quite slow, and only a thin surface covering was produced, which was removed during polishing.

The k/T relation for this material is given in fig. 2. It will be noticed that all curves peak at about 70°C. The D/E relation shows a definite saturation effect, both ends of the loop flattening towards the horizontal. The degree of saturation is less than that found in aluminium niobate (see below), but much greater than that found in manganese tantalate.

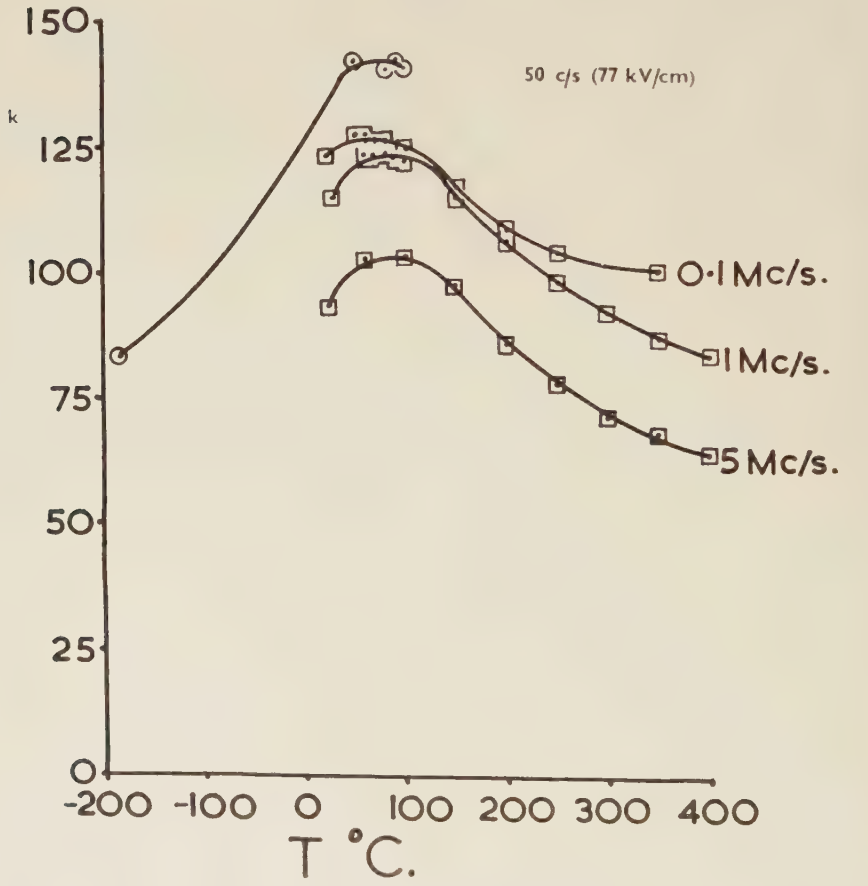
The k/T curves suggest the transition temperature is about 70°C. The peak position is difficult to observe at 50 c/s : above 70°C the increasing power-factor is causing an increase in the apparent value of k, thus obscuring the fall in true value. Saturation can still be observed in the D/E loop at 120°C, for an applied field of 77 kv/cm. It may be that at this temperature a sufficiently strong field will pull barium niobate into the ferroelectric state. However, the loss at 120°C for this ceramic specimen would be sufficient to mask the central linear portion if it exists, leaving only the saturation effects to show.

For frequencies up to 200 c/s the peak was very flat and ill-determined, as at 50 c/s. At 250 c/s the peak was clearly below 100°C, and at 1000 c/s was fixed as in the region of 70°C. It is thus likely that the position of the transition temperature is independent of the applied frequency.

The graphs of $1/k$ *vs* T are curved, showing that the k_0 contribution to the whole dielectric constant is large. A reasonable correlation can be obtained with the measured values of k and T at a frequency of 0·1 Mc/s for the following constants : $k_0 = 86$; $C = 6060°\text{K}$; $\theta = -82°\text{C}$.

Barium metaniobate is thus ferroelectric, with a transition temperature about 70°C.

Fig. 2



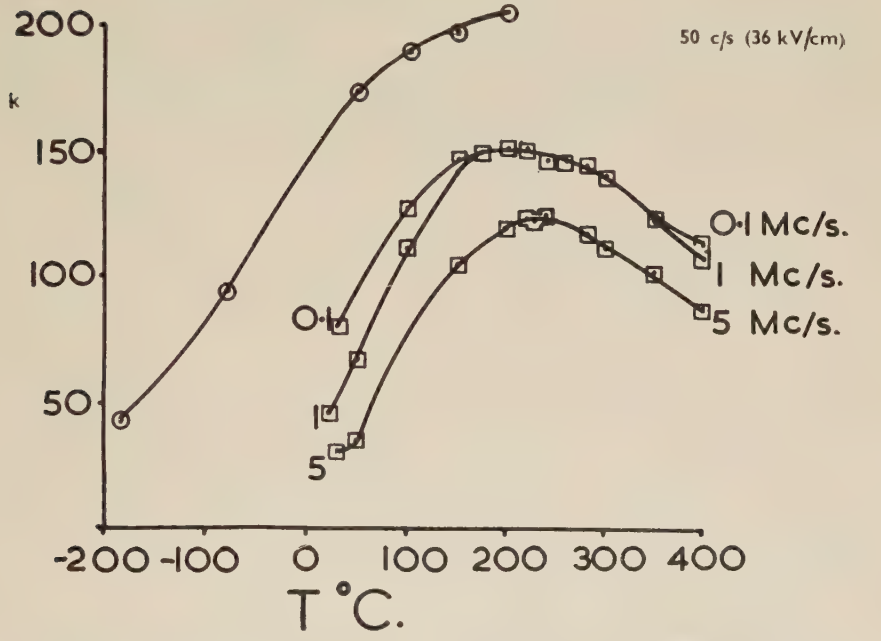
Relation between dielectric constant at different frequencies for $BaNb_2O_6$.

4.4. *Aluminium Metaniobate*, $AlNb_3O_9$

The 14 materials so far described have been of the general formula $A(BO_3)_2$, where A is a metal in a divalent state, and B is niobium or tantalum. Previous niobates and tantalates examined have been those

of the perovskite system $A(BO_3)_1$, where A is a monovalent metal ion. It would thus seem logical to continue the sequence to materials of general formula $A(BO_3)_3$, where A is now a trivalent metal ion. Ferric iron would be an obvious choice in view of the ferrous materials so far investigated, but the results obtained from ferrous niobate and tantalate suggested that the corresponding ferric compounds would be equally conducting, and hence give no definite information. The most stable trivalent metal readily available is aluminium, and therefore the niobate and tantalate of this were prepared.

Fig. 3



Relation between dielectric constant at different frequencies for $AlNb_2O_9$.

The D/E relation for the niobate was a thin loop curving gently towards the horizontal at each end and resembled that given by barium titanate ceramic when just below its Curie point. Due to increasing loss with increasing temperature no transition temperature could be found from these 50 c/s measurements.

The three r.f. curves, together with the 50 c/s curve, are given in fig. 3, and show quite clearly the transition to occur at about 220°c. For the r.f. values of k, graphs of $1/k$ *vs* T have been plotted. As the transition is at 220°c there are not many points above this temperature through which to draw the straight lines indicative of a Curie–Weiss law. However, the last four points of each set do lie on straight lines, and these have been

extrapolated to cut the temperature axis. Average values of the constants in the relation $k = k_0 + C/(T-\theta)$ for the lines are : $k_0 = 3$, $C = 50\,000°\text{K}$; $\theta = -80°\text{C}$.

Aluminium metaniobate is ferroelectric with a transition temperature about 220°C.

§ 5. Conclusion

Fifteen new metaniobates and metatantalates have been prepared and their dielectric properties examined in the ceramic form. Possibly three new ferroelectric materials—manganese metatantalate (very weakly), barium metaniobate, and aluminium metaniobate—have been discovered, and one new antiferroelectric lead metatantalate. Strontium and barium metatantalates may be of interest at temperatures below −183°C.

The research described here was carried out between October 1953 and September 1956. Since this latter date several Russian papers have appeared (e.g. Isupov 1957) dealing with metaniobates and metatantalates of divalent metals. Generally results are in good agreement with those given here, but there are two cases of severe disagreement.

(i) $PbTa_2O_6$ is reported as ferroelectric, with a transition temperature at 240°C ; $k = 310$–400 at 1 kc/s and 25°C.

(ii) $BaNb_2O_6$ is reported as non-ferroelectric, with $k = 40$ at 25°C and 1 kc/s, and having a negative k/T slope.

It must be emphasized that all results given in this paper are for materials prepared under the conditions quoted : polymorphism, with the various forms having quite different dielectric properties may well exist in several cases if production conditions are varied, as it does for $BaTiO_3$ and $PbNb_2O_6$. An independently produced supply of barium metaniobate ceramic discs gave properties and x-ray diffraction patterns exactly similar to those of the authors' own discs.

No electrical or optical results are available for single crystal specimens due to the difficulty of growth of large enough crystals. Such work is needed especially on barium metaniobate.

Little structural work has been done on any of these materials, again due to lack of single crystals and to the cell complexity. The iron and manganese compounds correspond to the mineral structures, all being the orthorhombic columbite cell except for iron tantalate which is the tetragonal trirutile dimorph. A complete structure analysis of any of these compounds is extremely difficult because of the low symmetry and large number of atoms per unit cell : Goodman (1953) describes lead niobate as having an orthorhombic cell of approximately 25 Å × 25 Å × 7 Å, with 40 molecules per cell. Francombe (1956) reports for the orthorhombic polymorph 17·51 Å × 18·81 Å × 7·73 Å, with 20 molecules per unit cell ; the large sides are both approximately $1/\sqrt{2}$ those measured by Goodman, thus giving a cell of half the volume.

Single crystals of barium metaniobate of size suitable for x-ray analysis were prepared. The unit cell is orthorhombic (space group either C_{2v}^{12}–$A2_1am$, or C_{2v}^{16}–$Ama2$), of size $10{\cdot}27\,\text{Å} \times 12{\cdot}16\,\text{Å} \times 7{\cdot}78\,\text{Å}$, having eight molecules per cell. A possible structure has been advanced from packing considerations (Coates 1956).

References

Coates, R. V., 1956, *Doctorate Thesis*, University of Bristol.
Cohn, G., 1948, *Chem. Rev.*, **42**, 527.
Cross, L. E., 1954, Research Correspondence, 7, p. S.36.
Francombe, M. H., 1956, *Acta Crystallogr.*, **9**, 683.
Goodman, G., 1953, *J. Amer. ceram. Soc.*, **36**, 368.
Isupov, V. A., 1957, *Izv. Akad. Nauk, S.S.S.R., Ser. fiz.*, **21**, 402.
Känzig, W., 1957, *Solid State Physics*, Vol. IV (New York : Academic Press).
Santen, Van J. H., 1946, *Trans. Faraday Soc.*, **42A**, 249.
Sawyer, C., and Tower, C., 1930, *Phys. Rev.*, **35**, 269.
Shirane, G., and Pepinsky, R., 1953, *Phys. Rev.*, **91**, 812.
Shirane, G., *et al.*, 1955, *Proc. I.R.E.*, **43**, 1738.

CORRESPONDENCE

Decay of Lithium-7 Hypernucleus

By P. H. Fowler
H. H. Wills Physical Laboratory, University of Bristol

[Received August 7, 1958]

Hyperfragments are so infrequently produced in nuclear reactions, that to the present time few cases of those with $Z \geqslant 3$ which offer an unambiguous interpretation have been found. This paper describes an event found in photographic emulsion which by good fortune can be interpreted with considerable confidence as the decay of ${}_{\Lambda}Li^{7}$. It gives a value of $5{\cdot}2 \pm 0{\cdot}5$ Mev for the binding energy of the Λ°. In addition, it is almost certain that the Λ° hyperon was produced directly from the capture of the K^{-}-meson by O^{16}.

The event was found in a stack of 600 μ Ilford K5 emulsion exposed to a separated beam of K^{-} particles at the Bevatron, Berkeley. A photomicrograph is given. The K^{-} adsorption at rest gave rise to three short black tracks (two of which are superimposed in the photograph, Pl. 103), one track of a lightly ionizing particle, and the hyperfragment, which decayed apparently at rest after travelling $32{\cdot}5\,\mu$. The hyperfragment decay resulted in a negative π-meson and two black tracks, which were coplanar to within the limits of error; the plane of decay was inclined at only 14° to the emulsion plane, so that accurate measurements could be made.

To obtain the best data on the angles of dip with respect to the emulsion plane, the dip measurements were made with a long working distance objective of numerical aperture equal to 0·95, and with the emulsion swollen with water so that it was approximately 1000 μ thick.

The table gives the data obtained for the hyperfragment decay.

Track	Range μ	Angle in common plane	Energy from range (Mev)	Momentum from range (Mev c)
π^{-}	10060	$0 \pm 0{\cdot}3^{\circ}$	$24{\cdot}1 \pm 0{\cdot}45$	$85{\cdot}5 \pm 1$
A He^{4}	33	$144{\cdot}5 \pm 1$	$6{\cdot}9 \pm 0{\cdot}1$	$227 \pm 1{\cdot}5$
B He^{3}	23	307 ± 1	$4{\cdot}9 \pm 0{\cdot}1$	165 ± 2

The directions of emission of the three tracks were coplanar to within the limits of error determined by multiple scattering and grain size; this suggests analysis as decay at rest. Using the measured range of the

π^--meson, one finds that tracks A and B should have momenta of 226 ± 16 and 164 ± 16 MeV/c respectively. If track A is due to He^4 its range should therefore lie between 27 and 39 μ or if He^3, 48 and 74 μ; any other assignment gives an even poorer fit. We conclude therefore that track A is very probably produced by an α-particle, and that its momentum is therefore $227 \pm 1{\cdot}5$ MeV/c. Using this estimate of momentum for A we obtained a new estimate for the momentum of track B, of 165 ± 2 MeV/c, so that if it is a He^3 nucleus, its range should lie between 22 and 24 μ, which it does; no other particle has a range close to 23 μ with this value of momentum. The whole event is thus consistent with the decay of a ${}_\Lambda Li^7$ hyperfragment as given below:

$$ {}_\Lambda Li^7 \rightarrow \pi^- + He^4 + He^3 + 35{\cdot}9 \pm 0{\cdot}5 \text{ MeV}. $$

The best estimate for the Q value for the reaction is obtained from the measured ranges, and the error is due almost entirely to the straggling of the π^--meson. This Q value gives a figure of $5{\cdot}2 \pm 0{\cdot}5$ MeV for the binding energy of the Λ° in ${}_\Lambda Li^7$, taking 37·2 MeV for the Q value for decay of a few free Λ°.

It is of interest to note that the other examples of ${}_\Lambda Li^7$ hyperfragments decay according to the schemes:

$$ (a) \quad {}_\Lambda Li^7 \rightarrow \pi^- + Be^7; $$

or

$$ (b) \quad \rightarrow \pi^- + H^1 + Li^6. $$

Identification of the first mode of decay is somewhat uncertain. For example, in the E.F.I.N.S.–N.U. collaboration (private communication) eight examples of collinear decays were found that could be attributed to this mode. The range of the recoil averaging 1·8 μ, however, does not permit an unambiguous assignment. A further important point here is that in an appreciable fraction of such decays, Be^7 might be expected to be produced in its first excited state at 430 kev; decay in this manner would not be resolvable from decay to the ground state, owing to straggling of the π-meson. A number of examples of the second mode of decay have been put forward (Ammar *et al.* 1958 and Limentani *et al.* 1958).

But there is difficulty in distinguishing this mode of decay from that of ${}_\Lambda Li^8 \rightarrow \pi^- + H^1 + Li^7$, the range of the recoil of Li^7 being $\sim 85\%$ of Li^6 for the same value of momentum—at values of momentum ~ 200 MeV/c, and an even greater fraction at lower values of momentum, where separation by range-momentum is not possible. There is no assistance to be obtained from the Q values of the two reactions. If any event gives an acceptable value for the energy release when analysed according to (*b*), it will give a value $\sim 0{\cdot}5$ MeV lower if the recoil is taken as Li^7 in a typical case; this will also be acceptable. In decays of type (*b*), once again the possibility of the Li^6 being produced in an excited state must be borne in mind.

Thus the event described in this paper is significant in that it is a clear example of the decay of ${}_\Lambda Li^7$, and it gives an accurate value for the binding

energy of the Λ°, free from doubt as to identification of the products and the possibility of the fragments being formed in any excited state. The value of $5{\cdot}2 \pm 0{\cdot}5$ MeV obtained is in fact in excellent agreement with the estimates from those decaying according to either reactions (*a*) and (*b*) and of an accuracy comparable to the average of all of them. Thus it is almost certain that the collinear decays have been correctly interpreted as decays of ${}_{\Lambda}Li^7$, but the question as to what fraction of the Be^7 recoils are produced in the ground state is left open.

The K^- Adsorption

Assuming the nature of the particles producing the three short black tracks it is possible to interpret the event as a K^- capture by O^{16} involving no neutral particles, and one obtains an excellent balance of energy and momentum.

The scheme proposed is:

$$K^- + O^{16} \rightarrow \pi^- + 2He^4 + H^1 + {}_{\Lambda}Li^7.$$

The resultant momentum of the four heavily ionizing tracks has a magnitude of 230 ± 10 MeV/c, and its direction lies very close to the observed light track; in fact the transverse momentum is only 8 ± 10 MeV/c. It is therefore very plausible that this track is that of the only other particle involved in the reaction, and that the reaction has been correctly interpreted. The final check that can be made is from the total energy release, which is 499 ± 9 MeV using this scheme, in excellent agreement with the K^- rest mass—494 MeV.

This excellent balance of momentum and energy obtained when one uses the reaction above, gives one confidence that it was very probably correct. Other schemes involving disintegrations of oxygen or heavy elements with emission of neutrons are unlikely to give such a good fit when analysed according to this reaction.

If one accepts this reaction, two further points can be made. First, we have further evidence of the correct identity of the hyperfragment, for no other hyperfragment having a momentum of 416 ± 6 MeV/c, as required from the dynamics of the parent star, has a range of $32{\cdot}5\ \mu$. Second, we can make a better estimate of the kinetic energy of the π, using the K^- mass; it is $124 \pm 0{\cdot}7$ MeV. This value is very high, and means that either the Λ° was made directly in the interaction of the K^- meson with a nucleon, or that if indeed a Σ hyperon was formed, it was formed in a bound state of ~ 30 MeV. The former supposition seems the more likely.

References

Ammar, R., Levi Setti, R., Limentani, S., Schlein, P. E., Slater, W. E., and Steinberg, P. H., 1958, *Proceedings Geneva Conference on High Energy Physics* (CERN). Enrico Fermi Institute for Nuclear Studies—North Western University, Collaboration.

Limentani, S. M., Schlein, P. E., Steinberg, P. H., and Roberts, J. H., Circularized Preprint, 1958.

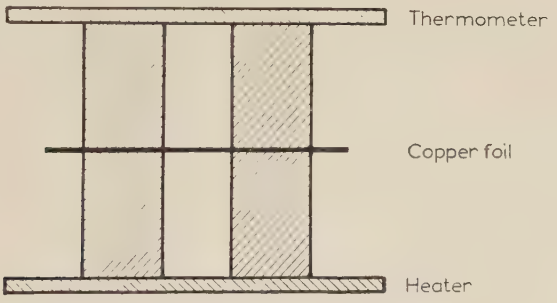
Heat Transfer between Copper and Liquid Helium II

By N. J. Brow† and D. V. Osborne
Department of Natural Philosophy, University of St. Andrews, Fife

[Received August 6, 1958]

When heat flows from a solid into liquid helium II, there exists at the interface a temperature discontinuity, the magnitude of which is proportional to the heat flow density. The constant of proportionality (usually expressed in deg watt^{-1} cm^2) has been called the boundary resistance or Kapitza resistance, K, after its discoverer (Kapitza 1941), and it has been measured subsequently by White *et al.* (1953), Fairbank and Wilks (1955) at temperatures below 1°K, and Dransfeld and Wilks (1957) at high pressures. All these measurements have been carried out with the aid of a constant heat current resulting in a steady temperature difference; we have thought it of interest to determine the same quantity using an alternating heat flow.

Fig. 1



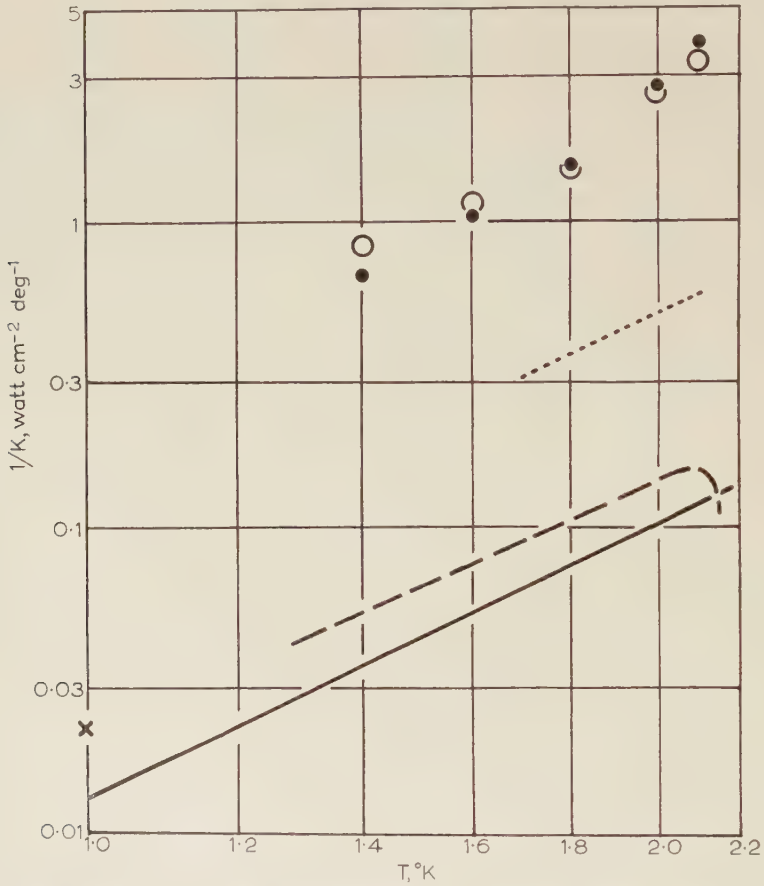
The coupled resonators.

The apparatus (see fig. 1) consists of two accurately similar second-sound resonators separated by a thin copper foil. Each resonator consists of a hollow cylinder of Perspex, 11 mm long and 6 mm inside diameter, closed at one end by a glass plate and at the other end by the copper foil. Second sound waves are generated in one resonator by means of an aquadag heater painted on the glass plate, and the second sound transmitted through the foil can be detected by an aquadag resistance thermometer painted on the other glass plate. The copper foil itself is sufficiently thin (0·05 mm) and of sufficiently high thermal diffusivity ($\sim 10^4$ cm^2 sec^{-1}) to

† Now at Chapel-cross Works, United Kingdom Atomic Energy Authority, Annan, Dumfriesshire.

offer no obstacle to the passage of thermal waves at the frequencies used (500 c/s to 2 kc/s). The Kapitza resistances on the two sides of the foil are, however, so great that the coupling between the two resonators is in fact relatively weak, and a measurement of the signal transmitted to the

Fig. 2



The dependence of $1/K$ on temperature; ● present work, fundamental; ○ present work, second harmonic; × Fairbank and Wilks (1955) extrapolated to 1°K; . . . . . Kapitza (1941); – – – – White *et al.* (1953); ———— theory of Khalatnikov (1952).

second resonator serves to determine the value of K. In a blank experiment, using a Perspex disc 9·5 mm thick in place of the copper foil, it has been established that the accidental coupling between the resonators is negligible compared with that due to the transmission of heat through the foil.

The determination of K in this way requires measurements of the phase and amplitude of the received signal as a function of frequency in the

vicinity of resonance. Since, moreover, it is not possible to measure absolute temperature amplitudes very reliably with aquadag thermometers, it is necessary to repeat the experiment with the same thermometer and heater but with the copper foil removed; such an experiment serves effectively to calibrate the thermometer at the frequencies used. The final results presented here have been derived from a series of four runs, two with the foil and two without, arranged so as to verify that the properties of the aquadag remained reasonably constant throughout.

The signal received in the second resonator depends not only upon the coupling, but also upon the attenuation present in each resonator and upon the difference in length, Δl, between the two resonators. Δl was small, about $5 \pm 5\,\mu$, and could not easily be measured accurately. The attenuation was therefore deliberately made appreciable by the use of a fairly narrow tube so that Δl became a negligible quantity in the final calculations.

The signals are small, and a phase-sensitive detector is therefore used. Such a device facilitates the measurement of phase, and at the same time has a very low noise level on account of its narrow bandwidth; a signal corresponding to $4\,\mu$deg can be detected. The amplitude has been kept sufficiently small to avoid non-linear effects; the received signal is in all cases sinusoidal and its amplitude is proportional to the input power. Measurements have been made for two different modes, the fundamental for which each resonator is half a wavelength, and the second harmonic for which each is a whole wavelength.

The results, together with those of previous workers, are shown in fig. 2. The accuracy of the present values of K is about $\pm 20\%$, and there is satisfactory agreement between values derived from measurements on the two different modes. The dependence upon temperature agrees with earlier work, but the actual magnitude of $1/K$ is somewhat higher. The theoretical curve of Khalatnikov (1952) is also shown for comparison. The disagreement between the earlier experiments themselves discourages us from laying too much stress on the absolute magnitude of any of the results, but if it should prove that the resistance measured with an alternating heat flow is genuinely different from the value obtained in a steady state experiment, the discrepancy may be associated with the turbulence present in a steady heat flow, but probably absent from second sound. This would imply that the mechanism of the Kapitza resistance is to be sought in the liquid itself as well as at the interface, and that the theory of Khalatnikov is therefore incomplete. The matter can be approached experimentally by studying second sound in the presence of a steady heat flow, though the apparatus would take on a more complicated form than that shown in fig. 1.

This work is published at this stage because it has had to be discontinued for the time being with no immediate prospect of its resumption. One of us (N.J.B.) is indebted to the Department of Scientific and Industrial Research for a grant during the tenure of which this work was carried out.

References

Dransfeld, K., and Wilks, J., 1957, *Proc. 5th Int. Conf. on Low Temp. Phys. and Chem., Madison.*
Fairbank, H. A., and Wilks, J., 1955, *Proc. roy. Soc.* A, **231,** 545.
Kapitza, P. L., 1941, *J. Phys., Moscow*, **4,** 181.
Khalatnikov, I. M., 1952, *J. exp. theor. Phys.*, **22,** 687.
White, D., Gonzales, O. D., and Johnson, H. L., 1953, *Phys. Rev.*, **89,** 593.

Electron Distribution in Transition Metals

By W. Hume-Rothery
Department of Metallurgy, Oxford University

P. J. Brown, J. B. Forsyth and W. H. Taylor
Crystallographic Laboratory, Cavendish Laboratory, Cambridge University

[Received September 19, 1958]

In a recent paper, Weiss and De Marco (1958) have used refined x-ray methods to determine the absolute scattering factors for crystals of transition metals of the First Long Period. The results are then analysed on the assumption that the scattering results from the sum of the effects of (1) the argon cores of the atoms concerned, (2) the electrons in 3d-like states, and (3) the electrons in 4s and 4p-like states. At the Bragg angles concerned, the effect of (3) is negligible, so that the number of 3d electrons can be obtained by subtracting the known effect of (1) from the experimentally determined structure factors. In this way, the following results are obtained for the numbers of 3d electrons per atom:

Cu	face-centred cubic	9·8
Ni	face-centred cubic	9·7
Co	hexagonal close-packed	8·4
Fe	body-centred cubic	2·3
Cr	body-centred cubic	0·2

In several recent theoretical papers, these numerical values have been taken as well-established experimental facts, and used to justify various bonding schemes or hypotheses. It is the object of the present note to suggest that great caution is needed before the results of Weiss and De Marco are used in this way, and that it is doubtful whether anything except a complete 3-dimensional electron–density map is of value for comparison with theoretical models.

The experimental work of Weiss and De Marco was done with great care and accuracy. The interpretation involving the subtraction of the effects of the argon cores is entirely reasonable, and may be regarded as justified by the results obtained for Cu whose Cu^+ ion contains 10 (3d) electrons and has a nearly spherically symmetrical electron cloud. For the remaining elements the interpretation is much less satisfactory, owing to the fact that only a few reflections of low indices were examined. The

critical sentence in Weiss and De Marco's paper is: "assuming that the residual scattering factors (after subtracting the 'argon core') for the lowest angle Bragg peak are entirely due to electrons with 3d-like radial charge densities, then that peak alone is sufficient to establish the number of 3d electrons". Actually, the measurement from a given reflection refers only to the corresponding direction in the crystal, and the method is, therefore, assuming that the electron density is spherically symmetrical. This is, however, the very thing which is in dispute, and on which the different theories differ so profoundly. It will be noted that, in the table given above, the apparent discontinuity between Fe and Co is also that between body-centred cubic metals (Cr, Fe) and close-packed metals (Co, Ni, Cu). The measurements refer to (110) reflections in the former, and to (100) (Co) and 111 (Ni, Cu) reflections for the latter, and it may be this difference which is revealed in the table given by Weiss and De Marco, rather than any discontinuity in the electronic configurations.

It should be emphasized that the elements Fe, Co, and Ni form a small group whose physical properties show such regular sequences that an abrupt break in the electronic structure between Fe and Co appears improbable. It may also be noted that, according to Deslattes (1958), the electronic structures of Weiss and De Marco are not in agreement with the details of the K-series x-ray spectra. The treatment of the 3d electrons as forming an independent group (and not as being in hybrid (4s 3d 4p) states) is of course equivalent to assuming what is in dispute between the different theories of the transition metals. It seems, therefore, that, although the experimental observations of Weiss and De Marco are accurate for the directions concerned, a complete three-dimensional electron-density map will be required before the electron distribution can be analysed with sufficient accuracy to be compared with the requirements of the different theories.

It should also be pointed out that Pauling (1953) was the first to put forward the view, which has recently been revived by later authors, that the saturation moment of $2{\cdot}2\,\mu_B$ for Fe results from atoms with the $(3d)^2$ atomic configuration, together with a slight contribution from interaction between the core and the conduction band.

Acknowledgments

One of us (W. H-R.) acknowledges with thanks the receipt of a grant from the British Non-Ferrous Metals Research Association towards the cost of a research programme of which the present work forms a part.

P. J. B. is indebted to the Department of Scientific and Industrial Research for the award of a Research Fellowship, and J. B. F. to Imperial Chemical Industries Ltd. for a research grant.

References

Deslattes, R. D., 1958, *Phys. Rev.*, **110**, 1471.
Pauling, L., 1953, *Proc. nat. Acad. Sci., Wash.*, **39**, 551.
Weiss, R. J., and De Marco, J. J., 1958, *Rev. mod. Phys.*, **30**, 59.

Nucleation of Cracks by the Intersection of Twins in α-Iron

By D. Hull
Metallurgy Division, Atomic Energy Research Establishment, Harwell

[Received August 8, 1958]

As part of a study of the nucleation of fracture in α-iron, the experiments described below have been made on a 3% silicon iron which had been produced by vacuum melting. Single crystal specimens 0·8 in. long and 0·15 in. wide were cut from a large-grained polycrystalline sheet, 0·002 in. thick, so that the cube edges of the body-centred cubic structure were parallel to the edges of the specimen. The cube face was in the plane of the specimen. Before straining, the specimens were electropolished.

The crystals were strained in tension along the [001] axis and in this way the maximum shear stress operated on four {011} slip planes. When tested at 78°K the crystals broke without measurable deformation taking place. Examination of the specimens after fracture at this temperature revealed that {112} twins had formed. No slip was observed. The final fracture consisted of {001} cleavage faces and irregular cracks along prominent twins. Cracks were not limited to the main fracture and as many as ten smaller cracks were observed in other parts of the specimen. It was clear that both the main fracture and the subsidiary cracks had formed at the intersection of twins. Two methods of nucleation were observed and these are illustrated in the photographs, figs. 1 and 2, Pl. 104. In the first case, fig. 1, the crack has formed along an {001} cleavage plane at the intersection two {112} twins; part of the crack has developed along one of the twins in a zig-zag manner, but retaining {001} cleavage facets. In the second, fig. 2, the crack has formed at the junction of one of the {112} twins with the other. In both cases the {001} cleavage crack has grown in both directions from the intersection.

Initiation of cracks by intersecting twins has previously been observed in body-centred cubic molybdenum (Cahn 1955) and in close-packed hexagonal zinc (Bell and Cahn 1958), but has not apparently been reported for α-iron. It has long been known that in tests at low temperatures on α-iron there is a connection between twinning and fracture, and two conflicting views have been taken, firstly that twins were nucleated by the high stress concentrations associated with fracture, and secondly that the formation of twins actually initiated the fracture as observed in the present experiments. It is probable that both effects occur. The mechanism by which intersecting twins produce a crack is as yet not known, but it may well be connected with Rose's channels (Rose 1868). Mügge (1922) estimated that there was a 50% increase in volume at the intersection of two twins on reciprocal planes, e.g. (112) and (112) as a result of Rose's channels in α-iron. He suggested that since this is unlikely cracks formed instead. However, Hall (1954) believes that the

strain set up is much more likely to be relieved by local slip instead of cracking. At very low temperatures and in structures with only a small number of slip systems, e.g. close-packed hexagonal, slip is difficult and such relief of strain will be less likely.

Barrett *et al.* (1937) reported that the stresses to initiate twinning and slip in iron and iron-silicon alloys depend on temperature and composition. In the material used for the present experiments the stress for twinning is less than that for slip at 78°K, and at this temperature the twins interact immediately to form cracks. It is hoped to make similar tests at higher temperatures, where slip is preferred to twinning, to determine whether cracks can be formed at the intersection of slip planes, as postulated recently by Cottrell (1958).

Acknowledgments

My thanks are due to Dr. A. H. Cottrell for his interest and encouragement and to Professor H. W. Paxton for his assistance.

References

Barrett, C. S., Ansel, G., and Mehl, R. F., 1937, *Trans. Amer. Soc. Metals*, **25**, 702.
Bell, R. L., and Cahn, R. W., 1958, *J. Inst. Metals*, **86**, 433.
Cahn, R. W., 1955, *J. Inst. Metals*, **83**, 493.
Cottrell, A. H., 1958, *Trans. Met. Soc. A.I.M.E.*, **212**, 192.
Hall, E. O., 1954, *Twinning* (London : Butterworths Scientific Publications).
Mügge, O., 1922, *Z. anorg. Chem.*, **121**, 68.
Rose, G., 1868, *Abh. preuss. Akad. Wiss.*, 57.

The Absorption of Sound in Dilute Solutions of Helium-3 in Liquid Helium II

By G. O. Harding and J. Wilks
Clarendon Laboratory, Oxford

[Received July 31, 1958]

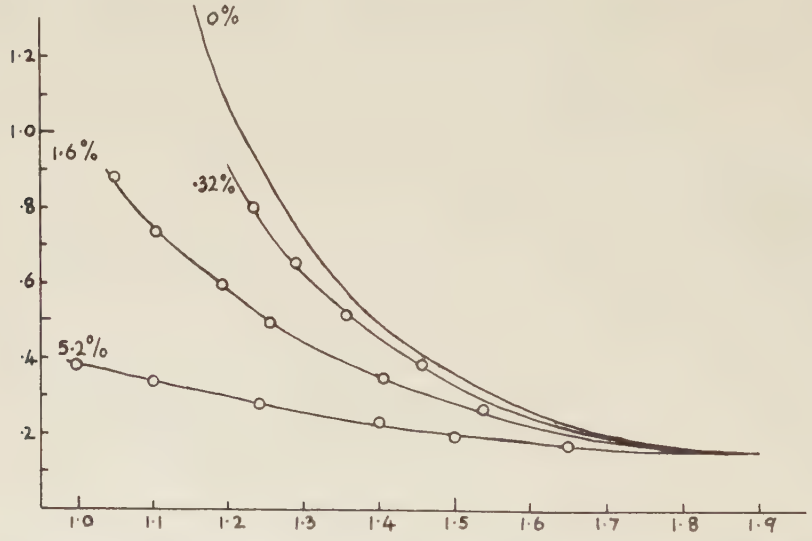
We have measured the coefficient of absorption of sound at a frequency of 14·7 Mc/s in liquid helium II containing small concentrations of ^{3}He. A pulse method with a variable path length was used as has been described previously (Dransfeld *et al.* 1958). The results for pure helium II and for concentrations of ^{3}He up to 5·2% are shown in the figure. It is obvious that below 1·8°K the presence of the ^{3}He considerably reduces the absorption.

It is convenient to discuss these results in the form of reduced absorptions equal to the observed absorption divided by the factor $2\pi^2 f^2/\rho c^3$. (The necessary values of the density ρ and velocity of sound c may be taken from the measurements of Dash and Taylor (1957) and Flicker and Atkins

(1957) respectively.) The absorption of sound in ^{3}He–^{4}He mixtures has been treated theoretically by Khalatnikov (1952a) who showed that for not too high frequencies the reduced absorption is given by

$$\alpha_r = \frac{4}{3}\eta_n + \zeta_2 + \frac{Dc^2}{\rho}\left(\frac{\partial\rho}{\partial\epsilon}\right)^2 \frac{m_3\epsilon}{kT}$$

where η_n and ζ_2 are the coefficients of first and second viscosity, D is a diffusion constant, m_3 the mass of a ^{3}He atom, and ϵ the mass concentration of the ^{3}He. The first two terms are similar in form to those obtained for pure helium II (Khalatnikov 1950, 1952b), while the last term is an additional one arising from the presence of the ^{3}He. Fluctuations in the density of the liquid set up by the passage of the sound wave produce concentration gradients of ^{3}He atoms, and diffusion then results in irreversibility and absorption.



The coefficient of absorption of sound of 14·7 Mc/s in liquid helium II containing ^{3}He. The figures on the curves indicate the molar fraction of ^{3}He.
Abscissae: Temperature °K.
Ordinate: Absorption (cm^{-1}).

The reduced absorption arising from diffusion can be calculated by estimating $\partial\rho/\partial\epsilon$ from the measurements of Dash and Taylor, and using the value of D found by Khalatnikov and Zharkov (1957) from the results of Beenakker *et al.* (1952); this contribution to the absorption is always small. The coefficient of first viscosity is also known from the experiments of Dash and Taylor; this term accounts for only 1 or 2 tenths of the total absorption. Thus the observed decrease in absorption can only be accounted for if the coefficient of second viscosity is considerably reduced by the presence of the ^{3}He atoms.

Second viscosity arises when the thermodynamic properties of a fluid do not immediately take their equilibrium values consequent on an

instantaneous change of pressure (see for example Wilks 1958). In liquid helium II, the numbers of thermal excitations do not immediately follow changes in pressure, and the approach to equilibrium is described by relaxation times characteristic of phonon–phonon and phonon–roton collisions. It appears that the presence of the ^{3}He atoms acts as a perturbation which increases the number of collisions which restore equilibrium, thus reducing the relaxation time and hence the absorption.

The variation of the absorption with concentration suggests that the number of additional collisions is proportional to the number of ^{3}He atoms present. Khalatnikov's treatment involves two relaxation times and his expression for the second viscosity is rather complicated, but for our purpose it is possible to introduce a simplification. Thus above 1·2°K the viscosity may be described tolerably well in terms of one relaxation time τ which varies with temperature in a similar way to the times given by Khalatnikov. From the general theory of relaxation processes

$$\zeta_2 \propto \frac{\omega^2\tau}{1+\omega^2\tau^2}$$

where ω is the angular frequency of the sound. In the present experiments, at all temperatures above 1·2°, $\omega \ll 1/\tau$ so that $\zeta_2 \propto \tau$ and $1/\zeta_2$ is proportional to the collision rate between the excitations. It follows that if $(\zeta_2)_0$ is the second viscosity in pure helium II and $(\zeta_2)_\epsilon$ is the viscosity for a concentration ϵ of ^{3}He atoms, then

$$[(1/\zeta_2)_\epsilon - (1/\zeta_2)_0] \propto N_\epsilon$$

where N_ϵ is the additional number of collisions due to the presence of the ^{3}He. An examination of the data shows that to a fair approximation

$$[(1/\zeta_2)_\epsilon - (1/\zeta_2)_0] \propto \epsilon$$

and one thus obtains the previously mentioned conclusion that the increase in collision rate is roughly proportional to the number of ^{3}He atoms. We are at present extending this work to lower temperatures.

Acknowledgments

We are grateful to Dr. K. Mayne for the mass spectrometer analyses of the ^{3}He concentrations.

References

Beenakker, J. J. M., Taconis, K. W., Lynton, E. A., Dokoupil, Z., and van Soest, G., 1952, *Physica*, **18**, 433.

Dash, J. G., and Taylor, R. D., 1957, *Phys. Rev.*, **107**, 1228.

Dransfeld, K., Newell, J. A., and Wilks, J., 1958, *Proc. roy. Soc.* A, **243**, 500.

Flicker, H., and Atkins, K. R., 1957, *Proc. Ohio Symposium on* ^{3}He.

Khalatnikov, I. M., 1950, *J. exp. theor. Phys.*, **20**, 243; 1952 a, *Ibid.*, **23**, 265; 1952 b, *Ibid.*, **23**, 8, 21.

Khalatnikov, I. M., and Zharkov, V. N., 1957, *J. exp. theor. Phys.*, **32**, 1108.

Wilks, J., 1958, *Z. phys. Chem.*, **16**, 372.

The Strength of Aluminium Silver Alloys

By Anthony Kelly
Department of Metallurgy, Northwestern University,
Evanston, Illinois

[Received September 26, 1958]

Single crystals of aluminium–6 atm % silver quenched from the δ-phase field and aged at room temperature show a critical resolved shear stress at 0°c of 7–8 kg mm^{-2}. Over the temperature range 77° to 373°k this value changes by not more than 10% (Lassila 1958). The value of the critical resolved shear stress is two orders of magnitude greater than that of pure aluminium single crystals. This increase in strength due to the addition of silver is intriguing since lattice strains are likely to be small in this system as the Goldschmidt radii of silver and aluminium differ by only 0·7% and solution of silver in aluminium produces no change in the x-ray lattice parameter (Axon and Hume-Rothery 1948).

Haessner and Schreiber (1956) have shown that amounts of silver less than the solubility limit (0·16 atm % at room temperature) produce no change in the plastic properties of aluminium single crystals. Lassila in this laboratory has found that aluminium specimens containing 6 atm % silver have an initial flow stress approaching that of pure aluminium, if tested below room temperature immediately after quenching from the solid solution range. 'Solid solution strengthening' thus seems to be small.

X-ray studies of these alloys (Walker and Guinier 1953, Gerold 1955) show that on ageing at room temperature the alloy contains zones consisting of spherical clusters of atoms containing a high concentration of silver surrounded by a spherical shell rich in aluminium atoms. A large volume fraction of the material is in the form of zones which are completely coherent with the matrix. Gerold finds that the hardness of the alloys depends only on the volume fraction of the zones and not on the zone radius or number of zones per unit volume. X-ray examination of the crystals studied here shows the cluster radius is *c.*17 Å and the radius of the zone *c.* 55 Å.

Although it cannot be proved conclusively that lattice strains play no part in determining the strength of these alloys it seems reasonable to look for alternative possibilities.

If we view the silver rich cluster and aluminium rich shell of each zone as forming a region in the alloy containing an equilibrium number of Al–Al, Al–Ag and Ag–Ag nearest neighbours, characteristic of a certain ageing temperature, the passage of a dislocation through a zone will alter the number of nearest neighbours of a given type across the slip plane. This will raise the free energy of the alloy and hence a higher stress will be

needed to move a dislocation through the alloy containing zones than through the random solid solution. This type of strengthening mechanism was suggested originally by Griffith (1920) and later by Fisher (1954) to account for the strength of alloys containing short range order.

Since the distance between the centres of zones, *c.*100 Å, is about twice their radius a dislocation cannot bow out between them but must pass through them approximately as a rigid line. The dependence of the strength solely on the volume fraction of the zones can then be explained in the following way. The average separation of the zones intersecting a given slip plane will vary as $(R^2/f)^{1/2}$ where R is the average radius of a zone and f the volume fraction of zones. The average length of the intersections of zones by a dislocation will vary as R and equating work done on the slip plane by the applied stress to the energy of the interface formed, a formula for the flow stress (σ) of the type

$$\sigma = \frac{f^{1/2}\gamma}{\alpha b} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

is obtained. Here γ is the energy of the interface produced when a zone is sheared by one interatomic distance. b is the Burgers vector and α is a constant which for spherical particles equals $2/\sqrt{\pi}$. Thus the flow stress is independent of zone size and hardening of this type should depend little on temperature at low temperatures, as is observed.

An estimate of γ is difficult. It will depend on the part of the zone through which a dislocation passes, on the width of the dislocation, and on the angle between a dislocation line and its Burgers vector. However, a rough value for a particular model can be obtained from the heat of reversion of the alloy (Kelly and Fine 1957). According to the x-ray results 10% of the total number of silver atoms in the alloy leave the clusters during reversion. Koster and Schell's (1952) value of the heat of reversion (1·04 cal. gm^{-1}) can then be expressed as 0·24 eV per atom of silver leaving the clusters. To a rough approximation we can write for the shearing of the clusters

$$\gamma = \frac{1}{3}\frac{\Delta E}{b^2}$$

where ΔE is the heat of reversion per atom of silver leaving the clusters. Whence $\gamma \sim 150$ ergs cm^{-2}. If we then calculate the flow stress according to eqn. (1) considering only the clusters at the centres of the zones and, again following the x-ray results, that the number of atoms in the clusters is one-third the total number of silver atoms we can take $f = 0{\cdot}3 \times 0{\cdot}06 = 0{\cdot}018$, and we find $\sigma \sim 7{\cdot}0$ kg mm^{-2}.

This estimate shows that this type of hardening should be of importance in these alloys and perhaps in other age-hardening systems where lattice strains are expected to be small, e.g. Fe–Cr and Cu–Co. Further, the flow stress should depend solely on the volume fraction of precipitate provided the spacing of precipitates is such that a dislocation passes through them as a rigid line.

Acknowledgment

This work was supported by the U.S. Air Force under Contract No. AF 18(600)–1468.

References

Axon, H. J., and Hume-Rothery, W., 1948, *Proc. roy. Soc.* A, **193,** 1.
Fisher, J. C., 1954, *Acta Met.*, **2,** 9.
Gerold, V., 1955, *Z. Metallk.*, **46,** 623.
Griffith, A. A., 1920, *Phil. Trans. roy. Soc.* A, **221,** 163.
Haessner, F., and Schreiber, D., 1956, *Z. Metallk.*, **48,** 263.
Kelly, A., and Fine, M. E., 1957, *Acta Met.*, **5,** 365.
Koster, W., and Schell, H. A., 1952, *Z. Metallk.*, **43,** 454.
Lassila, A. D., 1958, *M.S. Thesis*, Northwestern University.
Walker, C. B., and Guinier, A., 1953, *Acta Met.*, **1,** 568.

REVIEWS OF BOOKS

Logic Machines and Diagrams. By Martin Gardner. (McGraw-Hill Book Co., Inc.) [Pp. 152.] $4.75.

The syllogism dominated logic for two thousand years. The formidable weapon which finally brought about its downfall was symbolic logic, developed by Boole and others. Logical diagrams and machines also played their part by showing, in a non-mathematical and entertaining way, what the new logic was all about. That the subject can still be entertaining is shown in this book. Perhaps the machines described now strike us as complicated and creaking contrivances for what whey were able to do. However, this part of the subject is far from dead: enthusiasts still make logical machines, although they use relays of transistors instead of levers and wires. Formerly, problems that could be resolved by symbolic logic occurred only rarely in practice; now, such problems arise naturally in the design of complex switching circuits and it may be that a real demand for logical machines will be created. It is more likely, however, that any such demand will be met by the use of general purpose digital computers, suitably programmed, than by the construction of special machines.

The author of the book is described on the dust jacket as one of America's best-known writers in the field of science. It is not a popular book in the ordinary sense of the term; it would be just the thing, however, for readers of *The Philosophical Magazine* to take away on holiday. M. V. Wilkes.

Quantum Mechanics: non-relativistic theory. By L. D. Landau and E. M. Lifshitz. Translated from the Russian by J. B. Sykes and J. S. Bell. (London : Pergamon Press.) [Pp. 515.] 80*s*.

Amongst the numerous treatises on Quantum Theory, this is surely one of the most useful. It is not, apparently, meant as a first text. The reader would be supposed to know the elements of wave mechanics, and be used to thinking in terms of wave-packets and probability. But here he will find a detailed account of the subject (up to but not including the Dirac Equation) with elegant expositions of all the general principles, valuable theorems, and powerful techniques, many of which are only to be found in scattered papers and books. It is not clear whether the lucidity of the style stems from the Russian original, or from the translators; it could scarcely be bettered.

The tone is essentially pragmatic and eclectic. Purely mathematical questions (What axioms define Hilbert space ?) and philosophical doubts (Is quantum theory compatible with determinism ?) are entirely avoided. Von Neumann is only mentioned for a theorem on the intersection of molecular terms ! The subject is not forced into a restricted formal mould. Every technique is used where it is most powerful, so that the amount of mathematical manipulation is cut to a minimum. There is a happy balance between wave functions and matrix operators—between the Schrödinger and Heisenberg representations.

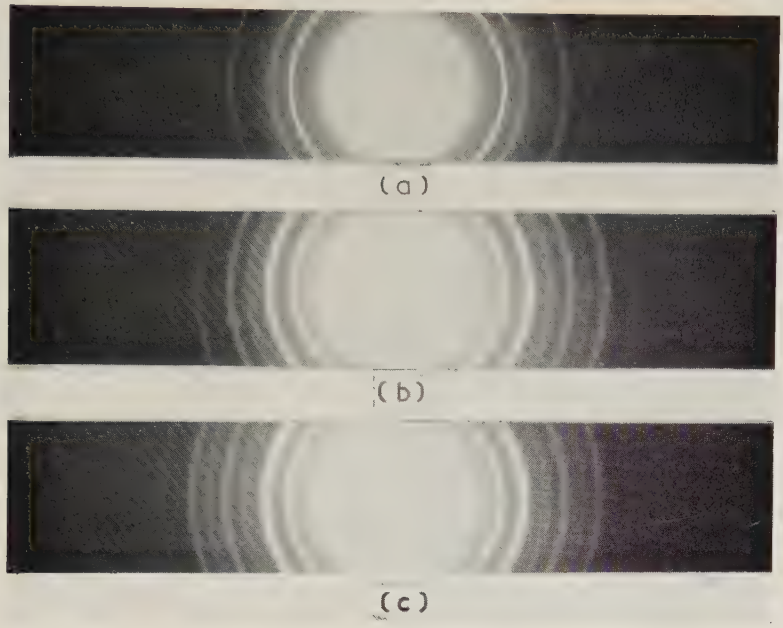
What is striking is the completeness with which the whole subject was developed between 1926 and, say, 1933. Almost all the references (which might perhaps have been given more chapter and verse than an author and a date) are from those heroic years. Indeed, this book, basically written in 1947, is a little old-fashioned in its choice of applications of the theory. We should now expect to include, along with the theory of atoms and molecules, the structure of the nucleus, and the electronic properties of solids. The many-body problem and the role of correlation and exchange energies, would also naturally demand more space. Perhaps these will be dealt with in other volumes of this admirable series by these distinguished authors. J. M. Z.

[*The Editors do not hold themselves responsible for the views expressed by their correspondents.*]

SUBJECT INDEX

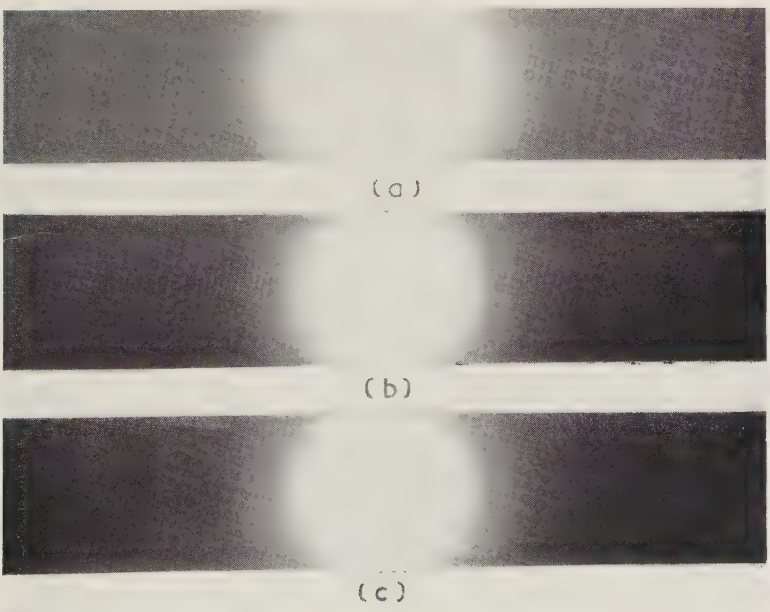
INDEX OF AUTHORS (WITH TITLES)

Fig. 6



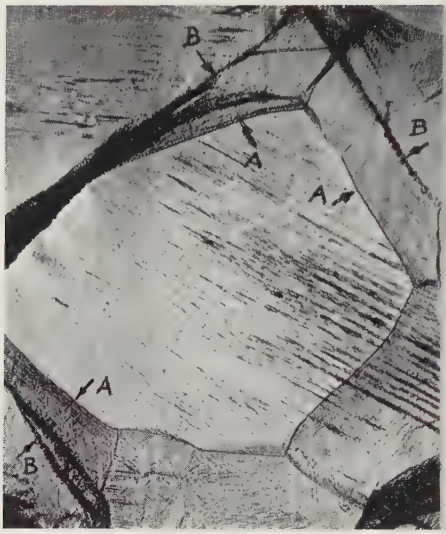
Electron diffraction patterns obtained from a specimen of aluminium–chromium by transmission. (*a*) Aluminium film 200 Å thick; (*b*) chromium film 150 Å thick; (*c*) compound film.

Fig. 7



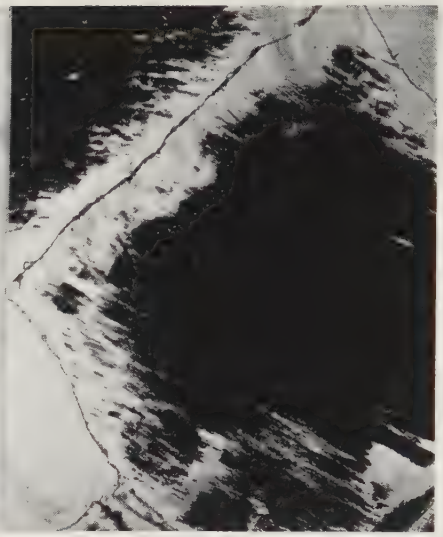
Electron diffraction pattern obtained from a very thin specimen of aluminium–chromium. (*a*) Aluminium film 60 Å thick; (*b*) chromium film 30 Å thick; (*c*) compound film.

Fig. 1



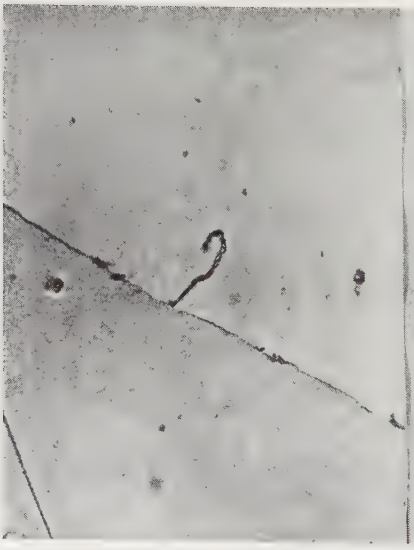
5×10^4 cycles ($\times75$).

Fig. 2



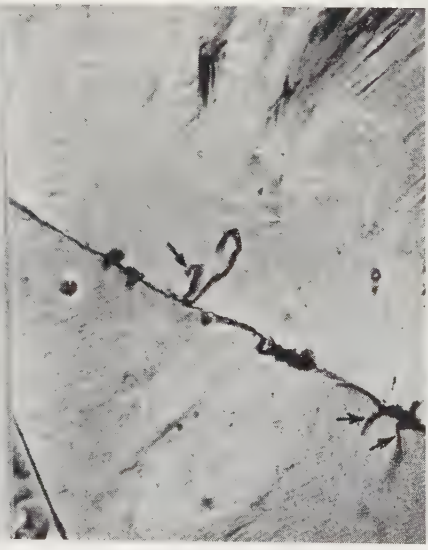
$2\cdot10^6$ cycles ($\cdot75$).

Fig. 3



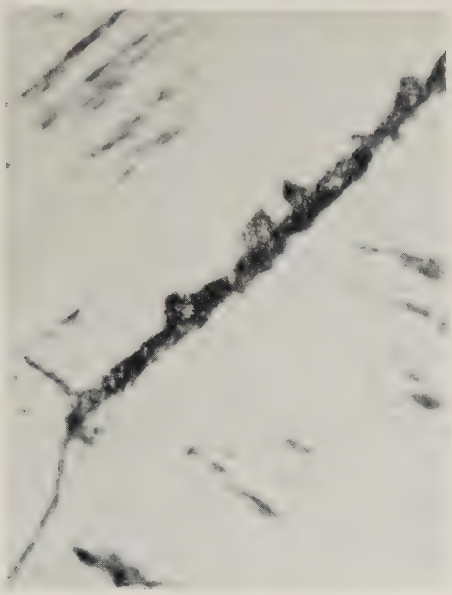
8×10^5 cycles ($\times354$).

Fig. 4



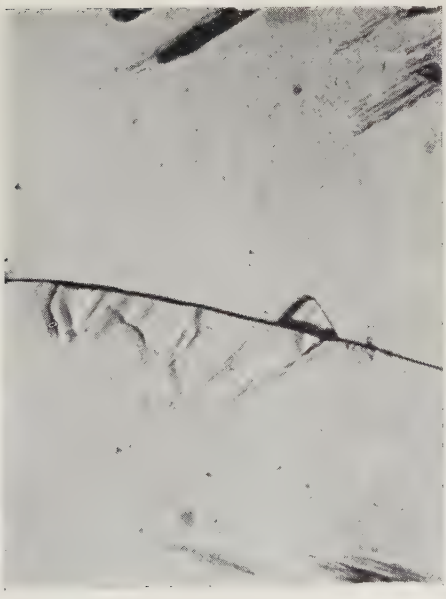
2×10^6 cycles ($\times354$).

Fig. 5



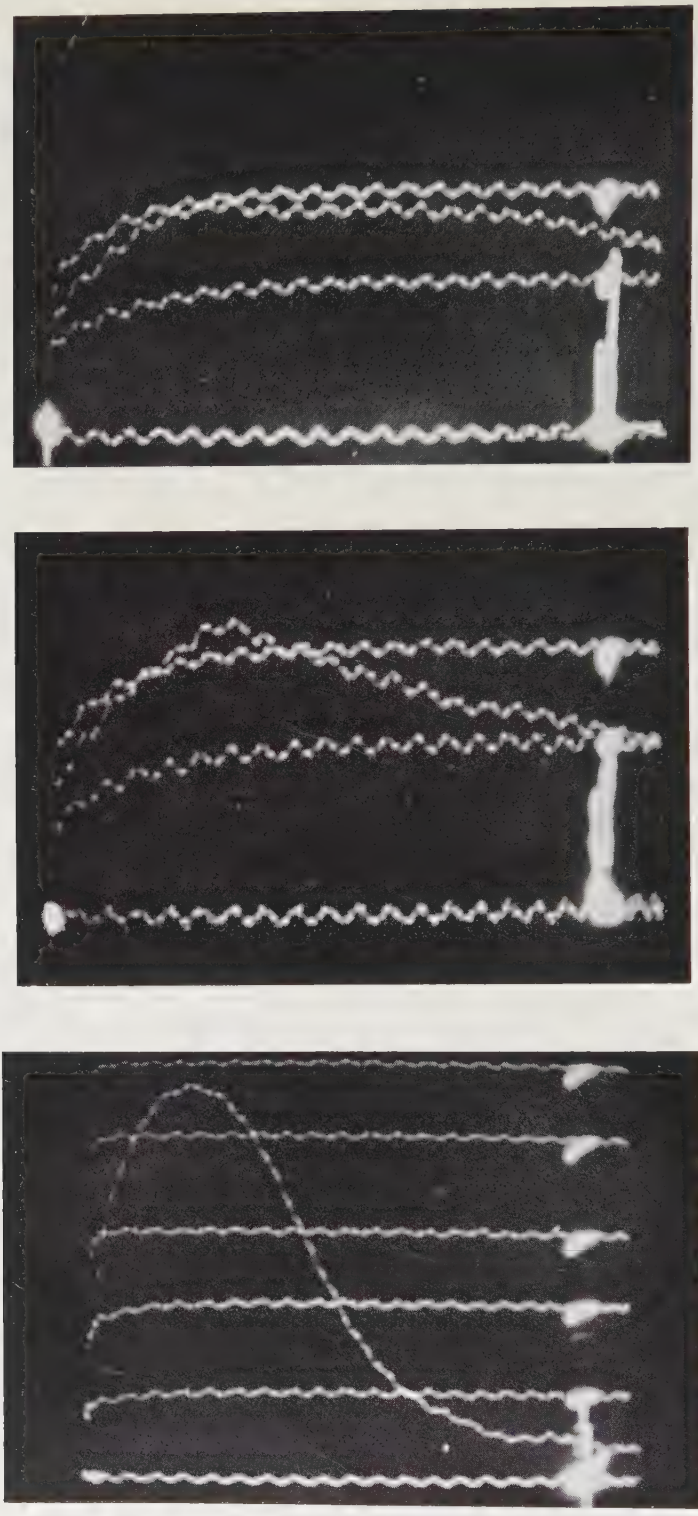
2×10^6 cycles ($\times 354$).

Fig. 6



2×10^6 cycles ($\times 354$).

Fig. 2



(*a*)

(*b*)

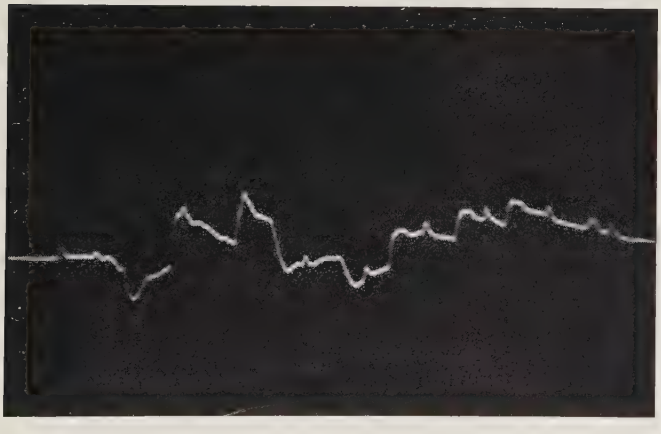
(*c*)

Strain-time oscillograms.

(*a*) Recorded from central gauge. Height of drop: 28 in.; Strains corresponding to calibration lines: 0, 0·92 and $1{\cdot}29 \times 10^{-3}$. (*b*) Recorded from series combination of central gauge and gauge mounted in a radial direction at 3 in. radius. Height of drop: 28 in.; Strains corresponding to calibration lines: 0, 0·92 and $1{\cdot}29 \times 10^{-3}$. (*c*) Recorded from gauge at centre of steel plate (0·330 in.). Impact velocity of steel sphere: 243 in./sec; Strains corresponding to calibration lines: 0, 0·34, 0·68, 0·97, 1·36 and $1{\cdot}70 \times 10^{-3}$.

Fig. 3

(*a*)



(*b*)

Typical oscilloscope records.

(*a*) Spiral oscilloscope record. Positive pulse blanking trace ; negative pulse brightening trace. One gyration takes 2 μsec.

(*b*) Type 58 oscilloscope record. Sweep time 1 μsec : 0·1 μsec calibration pulses. Eight events, polarity sequence $-++--+++$.

ORIGIN AND DECAY OF $_{\Lambda}Li^{7}$ HYPERNUCLEUS

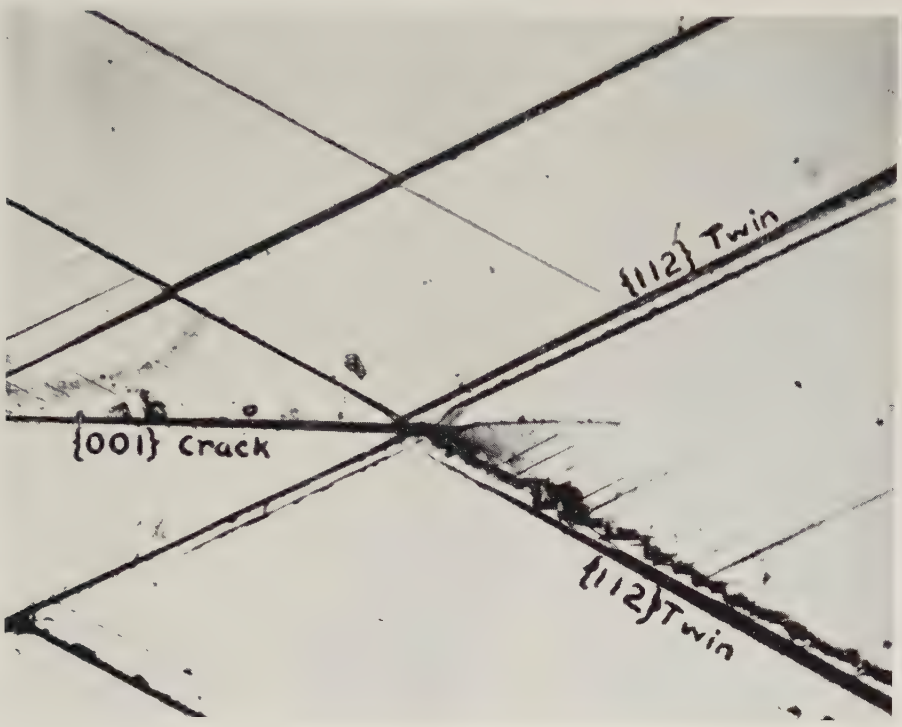
A negative K-meson is captured at rest by an oxygen nucleus, producing a fast π^{-}-meson, a $_{\Lambda}Li^{7}$ hypernucleus, a proton and two low-energy α-particles. The tracks of the two α-particles are superimposed and cannot be separately distinguished in the photograph.

The hyperfragment decays at rest according to the relation

$$_{\Lambda}Li^{7} \rightarrow \pi^{-} + He^{4} + He^{3},$$

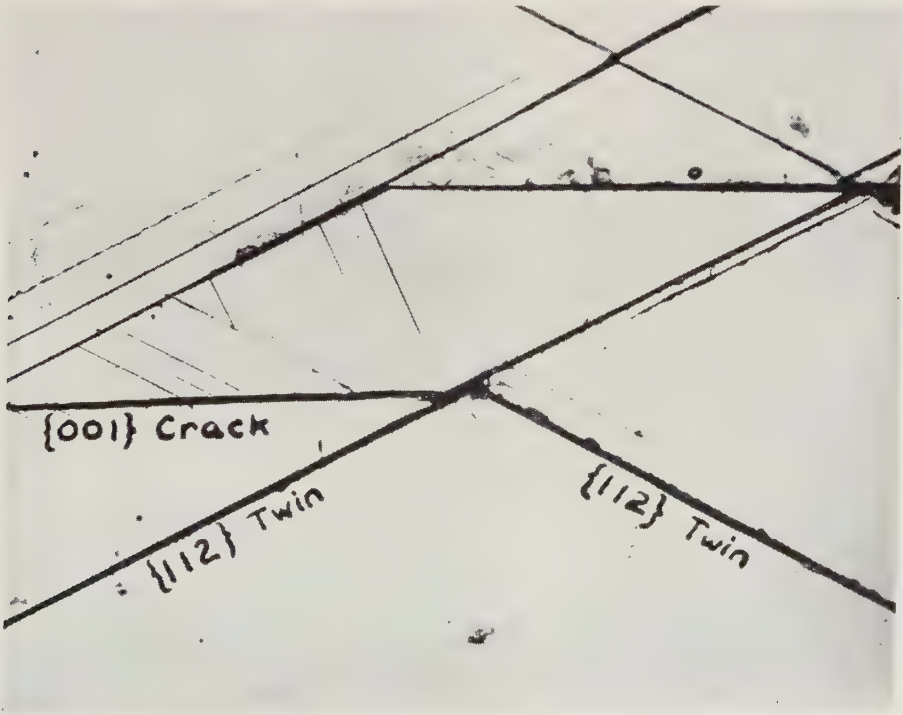
the He^{4} producing the longer of the two dense tracks.

Fig. 1



Crack nucleated by two intersecting twins. (×140.)

Fig. 2



Crack nucleated at the junction of one twin with another. (×140.)

END OF THE THIRD VOLUME

Printed by Taylor & Francis, Ltd
Red Lion Court, Fleet Street, London, E.C.4